



**Universidad Autónoma de Madrid  
Departamento de Medicina  
Facultad de Medicina**

**Measuring Disease  
Activity in Early  
Rheumatoid Arthritis:  
an emerging challenge**

---

**Cómo medir la actividad  
de la enfermedad en  
artritis reumatoide  
precoz: un nuevo reto**

**Isabel Castrejón Fernández, MD**

Madrid Junio 2014

---

*A José, por apoyarme siempre y embarcarse conmigo en  
el mayor proyecto de mi vida, Olivia*

**Universidad Autónoma de Madrid**  
**Departamento de Medicina**  
**Facultad de Medicina**

## **Measuring Disease Activity in Early Rheumatoid Arthritis: an emerging challenge**

### **Cómo medir la actividad de la enfermedad en artritis reumatoide precoz: un nuevo reto**

Tesis Doctoral presentada por la licenciada **D<sup>a</sup> Isabel Castrejón Fernández**, para optar al grado de Doctora en Medicina

#### **DIRECTORES:**

**Isidoro González-Álvaro, MD, PhD**

Jefe de Sección de Reumatología

Instituto de Investigación Sanitaria H. Universitario de la Princesa

**Loreto Carmona Ortells, MD, PhD**

Directora Científica del Instituto de Salud Músculo esquelética

**Rosario García de Vicuña Pinedo, MD, PhD**

Profesora Asociada Ciencias de la Salud (Departamento de Medicina, UAM)

Jefa de Servicio de Reumatología

Instituto de Investigación Sanitaria H. Universitario de la Princesa

## Contents

<b>ABSTRACT .....</b>	<b>5</b>
<b>RESUMEN .....</b>	<b>6</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>7</b>
<b>PART I – INTRODUCTION .....</b>	<b>8</b>
<b>Chapter 1: General Introduction .....</b>	<b>8</b>
Measuring disease activity in rheumatoid arthritis .....	8
Early and undifferentiated arthritis .....	11
Improving and defining new measure instruments .....	12
<b>Chapter 2: Objectives of the thesis.....</b>	<b>14</b>
<b>Capítulo 2: Objetivos de la tesis .....</b>	<b>15</b>
<b>Chapter 3: Methods .....</b>	<b>16</b>
General aspects .....	16
Patients and Procedures: The Princesa Early Arthritis Register Longitudinal study (PEARL).....	16
Disease activity assessment .....	18
Evaluation of the validation of available indices in early RA .....	19
Development and validation of a new index.....	20
Statistical procedures .....	22
<b>PART II – RESULTS.....</b>	<b>23</b>
<b>Chapter 1: Scopes and limitations of disease activity measures .....</b>	<b>24</b>
1.1. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? (Clin Exp Rheumatol 2008;26:769-75) .....	25
1.2. Estimated cut-off points for the 28-joint disease activity score based on C-reactive protein in a longitudinal register of early arthritis. (J Rheumatol 2010;37:1439-43).....	33
1.3. Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid. (Reumatol Clin 2010;6:134-40) .....	39
<b>Chapter 2: Recommendations for the investigation and follow-up of undifferentiated arthritis .....</b>	<b>47</b>
2.1. Clinical composite measures of disease activity for diagnosis and follow-up of undifferentiated peripheral inflammatory arthritis: a systematic review. (J Rheumatol Suppl 2011;87:48-53) .....	48
2.2. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. (Ann Rheum Dis 2011;70:15-24) .....	55

<b>Chapter 3: Development and validation of a new index .....</b>	<b>67</b>
3.1. Development and validation of a new disease activity index as a numerical sum of 4 variables in patients with early arthritis. (Arthritis Care Res (Hoboken) 2013 Apr;65(4):518-25).....	68
<b><u>PART III – GENERAL DISCUSSION</u>.....</b>	<b>77</b>
Summary of results.....	77
Strengths and limitations of the studies.....	80
Toward improved measures in early and undifferentiated arthritis .....	81
Treat to target in early and undifferentiated arthritis .....	82
Future perspectives .....	83
<b><u>PART IV: CONCLUSIONS</u> .....</b>	<b>85</b>
<b><u>PARTE IV: CONCLUSIONES</u> .....</b>	<b>86</b>
<b>REFERENCE LIST .....</b>	<b>87</b>
<b>ABBREVIATION LIST .....</b>	<b>92</b>
<b><u>PART V: ADDENDUM</u> .....</b>	<b>94</b>
Role of the funding sources.....	94
Curriculum Vitae.....	94
List of publications.....	95

## ABSTRACT

The objective of our research was to evaluate the use of composite indices to measure disease activity in rheumatoid arthritis (RA), with a special focus in early disease. It has been well established that early initiation of treatment improves patient outcomes. However, treatments used in patients with RA are not exempt of side effect; therefore, it is vital to adjust the treatment to each patient's disease activity.

This thesis is divided into a number of parts and chapters. The first part provides a general introduction about the options available for measuring disease activity in patients with RA, points to consider when addressing patients with an early disease, as well as the framework on how to improve or develop new instruments for measurement. In this opening section the aim and outline of the research and individual studies is also included, as well as the methodology followed to address the questions posed.

The second part contains the results of the different studies that support this thesis. A first chapter highlights some limitations of the available instruments; concretely, we address the study of such limitations in an early arthritis register, in which we propose new cut-offs for a version of the most widely used composite index, the DAS28 with CRP, given its limitations. The following chapter focuses on the choice of an appropriate instrument in patients with early disease, including the search for available instruments and recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis. A final chapter covers the development of a new composite index appropriate for patients with early disease. The third part of the thesis is a discussion of all the presented results pondered by the strength and limitations of the study approaches.

Finally, the conclusions illustrate the significance of improving measures in RA and future perspectives on this topic.

## RESUMEN

El objetivo de nuestra investigación fue evaluar el uso de índices compuestos para medir la actividad de la enfermedad en la artritis reumatoide (AR), con un enfoque especial en la enfermedad precoz. Ha sido claramente establecido cómo el tratamiento precoz mejora las medidas de desenlace del paciente. Sin embargo, los tratamientos que se utilizan en la AR no están exentos de efectos adversos por lo que es importante ajustar los tratamientos a la actividad de la enfermedad de cada paciente.

Esta tesis está dividida en distintas partes y capítulos. La primera parte consiste en una introducción general sobre las opciones disponibles para medir la actividad de la enfermedad en pacientes con AR, qué aspectos hay que considerar cuando se evalúan pacientes con enfermedad precoz, y una estrategia para mejorar o desarrollar nuevos instrumentos de medida. En esta parte inicial también se incluyen los objetivos, una descripción de la metodología de investigación y los estudios individuales así como la metodología seguida en cada pregunta de investigación.

La segunda parte incluye los resultados de los diferentes estudios que apoyan esta tesis. En el primer capítulo se destacan algunas de las limitaciones de los instrumentos disponibles. En concreto presentamos un estudio de la limitación de estos índices en un registro de artritis precoz, en el cual proponemos nuevos puntos de corte para el índice compuesto más empleado, el DAS28 calculado con PCR. El siguiente capítulo se centra en la elección del instrumento apropiado en pacientes con enfermedad precoz, e incluye una búsqueda de los instrumentos disponibles y recomendaciones sobre cómo investigar y hacer el seguimiento de los pacientes con artritis inflamatoria periférica indiferenciada. El capítulo final se centra en el desarrollo de un nuevo índice compuesto que sea apropiado para pacientes con enfermedad precoz.

La tercera parte de la tesis es una discusión de todos los resultados presentados haciendo hincapié en las fortalezas y limitaciones de los citados estudios.

Finalmente, las conclusiones ilustran el significado de mejorar las medidas en la AR y cuáles son las perspectivas futuras en este tema.

## ACKNOWLEDGEMENTS

I would like to thank all rheumatologists at Hospital Universitario La Princesa for their guidance and support during my years of fellowship, with special thanks to Dr. Isidoro González-Álvaro for initiating me as a researcher and Dr. Rosario García de Vicuña Pinedo for her enthusiastic support.

I wish to honor the memory of the late Dr. Armando Laffon, with gratitude for encouraging me to do my first rotation abroad at Hospital for Special Surgery (New York) in 2007. This rotation as an observer has been undoubtedly the first step in developing my career in the United States. In addition, it gave me the opportunity to meet Dr. Lilly Barillas who honors me with her unconditional friendship.

I thank Dr. Tom W. J. Huizinga and his colleagues for the warm welcome at the Department of Rheumatology at Leiden University Medical Center during my stay as a research fellow.

I want to thank all members of the 3E Initiative for offering me such a great opportunity to work as a team. In particular, I thank Dr. Loreto Carmona Ortell, not only for being my mentor in this project, but also for her help and support in countless projects and her valuable friendship. I also would like to thank Lucía Silva, Monica Schoels, Pedro Machado and Rosanne Koevoets. I have really enjoyed working with you all and I am very happy to keep such a wonderful friendship.

I am most grateful to Dr. Maxime Dougados for the opportunity he gave me to work in collaboration with the ESPOIR investigators.

Finally, I would like to thank Professor Ted Pincus for his guidance and support over the past four years, and for being not only my American Mentor but also a valuable example in my career as a researcher.



## **PART I – INTRODUCTION**

### **Chapter 1: General Introduction**

#### **Measuring disease activity in rheumatoid arthritis**

Rheumatoid arthritis (RA) is a potentially destructive disease with profound impact on patients' function and quality of life. It is also the most common chronic inflammatory joint disease affecting, in industrialized countries, 0.5 to 1.0% of adults, with 5-50 per 100.000 new cases each year (Carmona, Ballina et al. 2001; Carbonell, Cobo et al. 2008; Scott, Wolfe et al. 2010). Without any treatment, RA produces joint destruction with irreversible functional impairment, decreases quality of life and increases mortality (Pincus, Callahan et al. 1984). RA has a large range of articular and periarticular manifestations, including tenderness to palpation, morning stiffness, and motion impairment in the involved joints.

Over the last few decades, a better approach in the treatment of RA has resulted in considerable improvement in the outcome of patients. This has been the effect of multiple factors, namely the development of new therapies, such as the biological agents, a better use of classical therapies, and the use of strategies to frequently adjust the treatment according to disease activity, being the ultimate goal of therapy to maintain activity suppressed. The introduction and widespread use of quantitative measures, rather than the physician impression when making clinical decisions, has determined a notable advance in RA. Quantitative measures for RA are mainly composite indices, since a single measure—comparable to blood pressure in hypertension, or serum glucose in diabetes—cannot characterize disease activity status in all individual patients. In the clinical trial setting, composite indices make the comparison of the clinical efficacy of various treatments possible; in addition, these indices can be used to evaluate the effect of therapy in individual patients. They provide the additional bonus of offering a target at which treatment can be aimed, triggering adjustments as long as the target is not reached.

In 1993 the American College of Rheumatology defined a preliminary core set of disease activity measures for RA (Felson, Anderson et al. 1993). Most of the RA indices that have been developed are based on this core data set. The core data set includes three physician measures – tender joint count (TJC), swollen joint count (SJC), and physician's global assessment of

disease activity; three measures from the patient—assessment of pain, global assessment of disease activity, and of physical function; and a laboratory measurement of at least one acute-phase reactant, either erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP). Composite indices allow physicians to evaluate the indication and effect of particular therapies through an accurate assessment of disease activity. A correct choice use of these indices is especially critical to optimize therapy and patient outcomes. The best outcome for patients—namely, maintained low disease activity or remission—can be obtained by combining optimal treatment with appropriate assessment, switching therapies rapidly if the pre-established goal is not attained (Smolen, Sokka et al. 2003). As indicated by a European League Against Rheumatism (EULAR) taskforce on the management of RA, the choice of validated composite measures of disease activity and the target value may be slightly modified upon consideration of co-morbidities, patient factors, and drug-related risks (Smolen, Aletaha et al. 2010). The Disease Activity Score (DAS) was developed by van der Heijde and colleagues in the 1990s as a tool to measure disease activity in patients with RA (van der Heijde, van 't Hof et al. 1990; van der Heijde, van 't Hof et al. 1993). It combines, in a continuous score, the following components: the Ritchie articular index, this is the number of swollen joints (based on 44 joints), the ESR, and a patient's global assessments of disease activity (PGA) on a 0 to 10 cm visual analogue scale (VAS). The introduction of the DAS to measure disease activity in RA allowed comparing treatment effects between patients groups, as well as within individual patients. It has been proven to be a sensitive and specific tool to measure disease activity in RA (van der Heijde, van 't Hof et al. 1990). The level of disease control achieved when the DAS is used to monitor treatment response is greater than when more traditional endpoints are utilized (Grigor, Capell et al. 2004). The major downside of the DAS is that grading 44 joint counts may be time consuming. For this reason, a simplified version of the DAS that uses 28-joint counts, the DAS28, was then developed.

The DAS28 (Disease Activity Score 28) is the most widely used composite index in RA (Prevoo, van 't Hof et al. 1995). It includes a non-graded 28-joint count on both SJC and TJC. The joints assessed are the shoulders, elbows, wrists, metacarpophalangeal joints, proximal interphalangeal joints, and the knees (Figure 1). Rheumatologists use the DAS28 profusely and it is included in almost all clinical trials of RA.

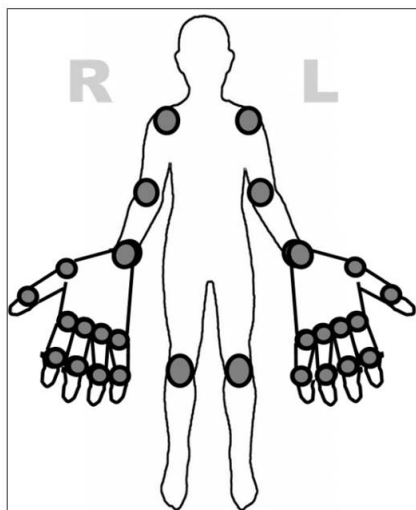


Figure 1: Joint counts based on 28 specific joints

However, the DAS28 has some problems. The main issue for clinical practice is that it is based on a complex equation and its determination requires a sophisticated calculator available on-line (<http://www.das-score.nl>), or as owner-developed Apps for mobile devices in the more recent years. Other concerns on the DAS28 have to do with its formula and metrics. For instance, in the DAS28 formula TJC is weighted twice as high as the SJC, when the latter is a more specific feature for RA than the former.

Bakker and colleagues also noted that, in the DAS28 formula, the ESR is heavily weighted and, therefore, it drives changes in the index even when ESR is within the normal range (Bakker,

Jacobs et al. 2007). Therefore, the DAS28 can lead to overestimation of disease activity in individual patients due to metric properties of its components.

A version of the DAS28 applies CRP instead of ESR in the formula (Inoue, Yamanaka et al. 2007; Matsui, Kuga et al. 2007), as in some patients; one of the two acute phase reactants may not be available. Unfortunately, the existence of two versions of the DAS28 may lead to both disagreements between indices and noise in measurement in clinical cohorts.

In an attempt to avoid the limitations of the DAS28, Smolen and colleagues developed years later the Simplified Disease Activity Index (SDAI) and a version without phase reactants, the Clinical Disease Activity Index (CDAI) (Smolen, Breedveld et al. 2003; Aletaha, Nell et al. 2005; Aletaha and Smolen 2005). These indices are an arithmetic sum of the variables included in the DAS28 including also a physician global assessment of disease activity (PhGA). SDAI includes CRP (mg/dl) as acute-phase reactant, whereas CDAI does not include any acute-phase reactant at all. These composite indices are derived as follows:  $SDAI = SJC28 + TJC28 + PhGA + PGA + CRP$  and  $CDAI = SJC28 + TJC28 + PhGA + PGA$ . These more simple measures were widely adopted for clinical practice, despite incomplete validation.

All of these indices provide an accurate assessment of disease activity in patients with high to moderate RA disease activity. However, when remission is addressed, SDAI and CDAI are more stringent than DAS28, the latter allowing 10 residual swollen joints in its classification of remission (Aletaha, Ward et al. 2005; Makinen, Kautiainen et al. 2005). This is particularly important to consider when deciding which composite index to use.

Both DAS28 and SDAI were developed and validated in populations with a definite diagnosis of RA, and only the DAS28 has been validated in populations of undifferentiated arthritis (Fransen, Visser et al. 2010). This is an important issue to consider since multiple studies have shown the importance of starting treatment as early as possible in RA, even before having a well-established diagnosis (Quinn, Conaghan et al. 2001; Nell, Machold et al. 2004).

Finally, despite the fact that adjusting treatment by a predefined target helps to suppress disease activity and to transform clinical remission into a realistic option, studies show that in daily practice rheumatologists do not work systematically towards achieving disease activity targets (van Hulst, Creemers et al. 2010). There might be some explanation for the low implementation of these indices in daily practice.

In summary, composite disease activity indices are important in providing a wide range of information and in allowing tighter control of therapy and supporting the optimization of treatment, but they exhibit some limitations to their applicability. For this reason, no recommendations exist on a universal single measure that should be applied in clinical practice. This situation offers the possibility to explore new potential indices with improved psychometric characteristics.

## Early and undifferentiated arthritis

Patients with RA, in general, present a typically distributed inflammatory polyarthritis of the hands, that by itself fulfils the American College of Rheumatology (ACR) classification criteria for the diagnosis of RA revised in 1987 (Arnett, Edworthy et al. 1988). However, some patients may present mono or oligoarthritis, showing a non-typical clinical picture that could fit any rheumatoid disorder, especially during the first weeks of the development of the disease. If no certain diagnosis can be established, the term “undifferentiated arthritis” (UA) is then used. In general, rheumatologists regard “early RA” patients as those presenting signs and symptoms of RA according to the ACR criteria for less than 3 months. However, the majority of patients are seen well beyond that time frame. For example, in UK the median total delay to see a rheumatologist is around 23 weeks (Kumar, Daley et al. 2007) and in Canada around 17 weeks (Feldman, Schieir et al. 2009). Early referral is improving because the urge for diagnosing and treating RA early, as a way to improve outcome, is well accepted by the rheumatology community (Aletaha, Eberl et al. 2004).

Recognizing RA as early as possible may improve its outcome and even prevent progression to chronic arthritis (van Nies, Krabben et al. 2013). There is a body of evidence that supports an earlier referral, as well as a more intensive treatment, and the widespread observance of these two maybe the reason why disease activity has become milder in recent years (Welsing, Fransen et al. 2005). However, identifying RA early poses a challenge. Ideally, biomarkers—including cytokines and genes—would be of great help. In fact, there is an growing body of research and theoretical development in this area. In a study by Raza and colleagues, patients with early arthritis, some of whom ended-up developing RA, were characterized by a distinct and transient cytokine profile (Raza, Falciani et al. 2005). This profile was no longer present in established RA, indicating that the initial phase of RA may be governed by distinct pathophysiologic mechanisms. Interventions targeting these early events could prevent progression to RA. Nevertheless, the greatest impediments for the use of biomarkers are their limited diagnostic performance, high variability among patients, and costs.

The first RA biomarker was rheumatoid factor (RF), discovered in 1948 (Rose, Ragan et al. 1948). The finding of this complex immunoglobulin molecule raised hope that diagnosis and management of RA would be better from then on. It is well known that this biomarker is typical of RA patients, and in fact it became part of the 1987 RA criteria, and it is also associated with a more severe disease (Mewar, Coote et al. 2006). A variety of other autoantibodies was described in more recent years, with a remarkable interest in the antibodies to citrullinated peptides (anti-CCP or ACPA). Anti-CCP show higher specificity than RF to identify patients with RA (Nishimura, Sugiyama et al. 2007), although in general both biomarkers overlap widely (Nell, Machold et al. 2005). Both, anti-CCP and RF, are now included in the 2010 RA classification criteria (Neogi, Aletaha et al. 2010)

Biomarkers could also be helpful to predict response to therapy; however, up to date there are no strong predictive markers for remission. In contrast, some baseline independent variables, not properly biomarkers, predict remission much more strongly, such as low disease activity

(DAS<4) and low health assessment questionnaire (HAQ) score (<1.25)(Gossec, Dougados et al. 2004).

Imaging could also play an important role to evaluate patients at an early stage of the disease. Joint destruction is generally assessed by traditional radiographs, but radiographic changes are presented in a later stage of the disease. Other imaging modalities, such as magnetic resonance imaging (MRI) and ultrasonography, have been pointed out to improve detection of erosive disease in early phases of RA. But until now, neither MRI nor ultrasonography images have shown a sound capacity for predicting radiographic changes typical of RA. In a study, which evaluated patients with no erosive polyarthritis in hands and different diagnoses, the arthritis images present in RA patients were no different than MRI images from patients with systemic lupus erythematosus or Sjögren's syndrome, with the only exception of a higher frequency of bone marrow oedema in the metacarpophalangeal joints of patients with RA (Boutry, Hachulla et al. 2005).

In summary, early control of disease activity and intensification of therapy according to the level of disease activity lead to better outcomes compared to a traditional approach. However, evaluating a patient at this stage of the disease is complicated. Neither biomarkers nor imaging techniques have shown to be the ideal approach to evaluate these patients. The use of composite indices, which summarize the clinical picture, can provide a useful way to evaluate patients with early disease. But it is important to consider that the majority of composite indices have not been validated in early arthritis.

## Improving and defining new measure instruments

Therapy for RA has seen great progress over the past years, including the approval of new drugs and the implementation of new strategies. The use of measure instruments has contributed undeniably to a better management of the disease, but these measures need to be reliable and valid, not only in clinical trials but also in the clinical setting, as a requisite to be fit for use in the follow-up of patients.

The DAS28 has been of great use to the rheumatology community for more than a decade. The simpler formulae of the SDAI and CDAI have shown similar validity, with improved feasibility. However, all these composite indices have limitations. In addition, and despite growing evidence showing that standardizing disease activity assessment and treating to target are effective (Schoels, Knevel et al. 2010), only 15.2% of rheumatologists use CDAI and 27.8% DAS28 routinely in clinical practice (Anderson, Caplan et al. 2012); furthermore, in randomized clinical trials (RCT), measures are used with undesirable heterogeneity.

Very plausibly, the large availability of instruments and the lack of agreement on which indices are the best to measure disease activity, may contribute to their inadequate implementation in clinical practice and RCT. In an effort to provide some guideline, the ACR has recently performed a review to comprehensively evaluate the validity of available RA disease activity measures (Anderson, Caplan et al. 2012). In this review, 63 currently available RA measures were identified and, in a multistep process, 6 composite indices were established as the

recommended ones: CDAI, DAS28, PAS, PAS-II, RAPID3, and SDAI. Nevertheless, the authors recognize that there is no ideal measure, and that all of them showed a similar level of adequacy and psychometric properties.

Since the development of the first composite disease activity measure for use in RA, many attempts have been made to improve RA disease activity monitoring. There is always a need to improve our measuring tools, especially in the early stages of the disease. The introduction of new tools is aimed to create more valid indices, both from the physician and the patient perspectives. In this sense, it is important to assure that each new composite index has comparable validity, so that physicians or researchers can pick the tool that works better in their clinical or research setting.

## Chapter 2: Objectives of the thesis

The general objective of the thesis is to gain insight in the measurement of RA, concretely in the measurement of early arthritis.

For these, the following objectives were established:

1. To explore potential biases that may contribute to the observed variability in the DAS28 score.
2. To analyze the variability in the DAS28 response that could be specifically attributed to a gender bias.
3. To estimate the best cut-off points for the DAS28-CRP to classify patients with early arthritis.
4. To analyse the individual behaviour of each component of the DAS28 in early arthritis patients.
5. To identify measures properly validated in undifferentiated arthritis.
6. To recommend measures for undifferentiated arthritis based on an analysis of their validation.
7. To identify gaps in the validation of measures in undifferentiated arthritis.
8. To develop a new composite index appropriate to use in patients with early disease.

## Capítulo 2: Objetivos de la tesis

El objetivo general de esta tesis es adquirir conocimiento en la medida de la AR, en concreto en la medida de artritis precoz.

Para ello, se han establecido los siguientes objetivos:

1. Explorar los sesgos potenciales que pueden contribuir a la variabilidad observada en el DAS28.
2. Analizar la variabilidad en la respuesta del DAS28 que podría ser específicamente atribuida a los sesgos de género.
3. Estimar los mejores puntos de corte para el DAS28 calculado con PCR que permitan clasificar pacientes con artritis precoz.
4. Analizar el comportamiento individual de cada componente del DAS28 en pacientes con artritis precoz.
5. Identificar medidas validadas adecuadamente en pacientes con artritis indiferenciada.
6. Recomendar medidas para artritis indiferenciada de acuerdo a su validación.
7. Identificar falta de datos en la validación de medidas en artritis indiferenciada.
8. Desarrollar un nuevo índice compuesto apropiado para pacientes con enfermedad precoz.



## Chapter 3: Methods

Although the methods of each study included in this dissertation are addressed at each of the manuscripts that conform this thesis, general aspects of the methodology employed, as well as clarifications and definitions are presented in this chapter.

### General aspects

In the initial part of this thesis, we explored the potential biases of the use of ESR or CRP indistinctly to calculate DAS28. In a later approach, we describe a set of new cut-offs for the DAS28 version including CRP. Finally, in a third part, we focus on the influence of gender in the evaluation of disease activity. For these three objectives, we used data from the Princesa Early Arthritis Register Longitudinal study (PEARL) established in La Princesa Hospital in September 2000.

The second part of this thesis evaluates which composite measure may be the most appropriate instrument to evaluate patients with early disease. As such, this part acts also as a summary of recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis. For these objectives, systematic literature review and consensus methodology was used.

The final part covers the development of a new composite index appropriate to use in patients with early disease. Each aspect of the validation of this new index is explored in the previously mentioned PEARL cohort.

### Patients and Procedures: The Princesa Early Arthritis Register Longitudinal study (PEARL)

Patients for the analyses included in the first and the last part of this thesis, were selected from a longitudinal register established in La Princesa Hospital from September 2000, the Princesa Early Arthritis Register Longitudinal study (PEARL). The objective of this register is to study the clinical course, prognostic factors and biomarkers, as well as response to treatments in early

arthritis patients. This is a single-centre clinical and on-going register of patients attending a dedicated Early Arthritis Clinic (EAC). The area of influence covers a population of 500,000 inhabitants, of whom more than 90% are covered by public health insurance. In addition, all primary care physicians in the area are aware of the EAC and of the procedures to refer patients to it.

To be referred to the EAC, patients have to have had two or more swollen joints for at least four weeks and symptoms for less than a year. Patients are subsequently excluded from the study, and referred to standard rheumatology care, if they are diagnosed of gouty arthritis, septic arthritis, spondyloarthritis, or connective tissue diseases during follow-up. All patients attending the EAC sign a written consent form upon first visit, after being informed about the details of the procedures, which include data collection for research studies. The institutional research board and ethics committee of the hospital approved the procedures of the clinic for research.

The protocol includes a baseline visit and three visits during a follow-up period of two years. Subsequently, protocol visits are scheduled annually. At each visit, the following information is collected per protocol and entered into an electronic database: 1) clinical and demographic information; 2) data on treatment, including disease-modifying antirheumatic drugs (DMARDs), steroids, and nonsteroidal anti-inflammatory drugs (NSAIDs); and 3) RA clinical measures, such as 28 TJC and SJC, global disease activity on a 100 mm visual analogue scale assessed both by the patient and the physician, and basic laboratory tests, including ESR and CRP. Two experienced rheumatologists perform all joint counts in order to reduce inter-rater variability. Conventional radiographs of hands and wrists are performed at baseline and then annually at follow-up.

At each visit, the date of initiation and discontinuation of DMARDs, as well as the maximum and minimum doses reached through the follow-up, are systematically collected. The protocol does not contemplate a specific treatment strategy. Treatment is prescribed according to the clinical judgment of the rheumatologist, with as many visits as deemed necessary by the treating physician between the protocol data collection points.

Around 62% of the patients followed-up in PEARL fulfil the ACR criteria for the diagnosis of RA (Arnett, Edworthy et al. 1988), and 38% have undifferentiated arthritis, at their first visit. The PEARL patients are similar to those patients from the Leiden Early Arthritis Clinic, providing support for the generalizability of the PEARL population (Table 1).

**Table 1:** Comparison of the Leiden Early Arthritis Clinic and the PEARL.

	L-EAC			PEARL		
	Women N=160	Men N=57	<i>p</i>	Women N=150	Men N=45	<i>p</i>
Age at inclusion	55.9 (14.9)	60.8 (13.5)	<b>0.030*</b>	51.1 (16)	59.7 (15)	<b>0.002*</b>
DAS44/DAS28	3.51 (1.01)	3.32 (0.91)	0.36	4.71 (1.46)	3.76 (15)	<b>&lt;0.001*</b>
Tender Joints	10 (6)	8 (6)	0.21	7 (7)	4 (4)	<b>0.002*</b>
Swollen Joints	8 (7)	9 (6)	0.43	5 (5)	5 (5)	0.47
VAS general	44.6 (25.3)	35.1 (25.5)	<b>0.023*</b>	45.2 (23.2)	37.9 (19.4)	<b>0.039*</b>
VAS pain	52.7 (23.2)	44.9 (22.4)	<b>0.032*</b>	49.2 (26.2)	35.8 (22.7)	<b>0.003*</b>
HAQ	1.18 (0.72)	0.99 (0.66)	<b>0.014*</b>	1.17 (0.73)	0.75 (0.64)	<b>&lt;0.001*</b>
ESR	35.8 (24.4)	34.4 (23.6)	0.71	30.9 (21.5)	24 (19.2)	0.07
CRP	24.4 (33.3)	30.2 (30)	<b>0.04*</b>	14.8 (27.1)	27.3 (48.7)	<b>0.04*</b>
FR+	101 (65.2%)	35 (61.4%)	0.61	67 (44.7%)	17 (37.8%)	0.41
Anti-CCP+	53 (56.4%)	19 (55.9%)	0.96	63 (41.9%)	12 (25.6%)	0.053

\*Statistically significant comparison of groups within cohort. Data are presented as mean values (standard deviation), unless otherwise specified. Tender and swollen joint counts are based on 44 joints in the L-EAC and 28 in the PEARL.

## Disease activity assessment

Composite indexes are calculated following the recommendations provided by each developer. Regarding DAS and DAS28, the formula provided in the official website [<http://www.das-score.nl>] was applied to calculate the score. Minimal disease activity (MDA) is calculated as described by Wells et al. (Wells, Boers et al. 2005), being equal to 1 when the patient meets 5 of the following criteria: (1) Pain (0–10) ≤ 2; (2) SJC (0–28) ≤ 1; (3) TJC (0–28) ≤ 1; (4) HAQ (0–3) ≤ 0.5; (5) PhGA (0–10) ≤ 1.5; (6) PGA (0–10) ≤ 2; (7) ESR ≤ 20.

To generate the new set of cut-off points for DAS28-CRP, a retrospective evaluation of disease activity by six rheumatologists was used as “gold standard”. They were four senior rheumatologists and two fellows from La Princesa Rheumatology department. Half of them were women, and they had a median experience in assessing RA patients of 13.2 years. Disease activity was graded as remission, low activity, moderate activity or high activity at each visit, based on the following information: TJC, SJC, PGA, HAQ, ESR and CRP. They performed no physical examination of the patients. In addition, they were blind to the PhGA obtained at each

visit. There was total agreement between these 6 rheumatologists in 84% of the visits; in the remaining 16%, disease activity status was established by consensus.

## **Evaluation of the validation of available indices in early RA**

Very timely, two of us became part of an international project, the 3E (Evidence, Expertise and Exchange) initiative 2008-9, to develop recommendations on how to investigate and to follow-up undifferentiated peripheral inflammatory arthritis (UPIA). Our task was to perform one of the 10 systematic literature reviews and to present it to a broad international panel of rheumatologists. In total, 697 rheumatologists from 17 countries participated in this initiative, each country represented by a scientific committee consisting of one principal investigator and 5–13 members. Ten international fellows performed the systematic reviews supervised by five mentors.

During the first international meeting of the project, 10 clinically relevant questions related to UPIA were formulated and selected via a modified Delphi vote. The relevant question that was assigned to us was which measures of clinical disease activity should be used in UPIA.

The clinical question was approached using the PIO format (Patients; Intervention; Outcomes or target conditions). The definition of UPIA is controversial and there is no widely accepted classification criterion for this condition. During the 2008–9 3E Initiative kick-off meeting, experts decided that only patients in whom clinically apparent joint swelling (synovial proliferation or synovial effusion) was observed by the rheumatologist should be included. For this review we systematically searched for studies of patients who did not fulfil diagnostic or classification criteria for any specific rheumatic disorder at baseline. Studies with mixed populations (e.g., including arthralgia or early RA) were also retained.

A systematic literature search for articles published up to February 2009 was carried out in Medline, Embase, and Cochrane Library using comprehensive search strategies elaborated in collaboration with experienced librarians. Retrieved citations were screened for titles, abstracts and full text using predefined inclusion and exclusion criteria; full read papers and review articles were hand-searched for additional references. Retained articles were graded for their methodological quality according to the levels of evidence of the Oxford Centre for Evidence-Based Medicine. For each question, relevant data were extracted and appropriate statistics were calculated. In our review, we analysed the level of validation of the measures applying the definitions presented in the next heading.

A national meeting was held in each country to discuss the generated evidence and to propose a set of recommendations. In a third joint meeting the 17 scientific committees merged all propositions into 10 final recommendations via discussion and modified Delphi vote. The grade of recommendation according to the Oxford levels of evidence was attributed and the level of agreement was measured on a 10-point numerical rating scale (1=no agreement, 10=full agreement). Finally, the potential effect of each recommendation in clinical practice was assessed according to three impact statements voted by the rheumatologists.

## Development and validation of a new index

The Hospital Universitario de la Princesa Index (HUPI) was developed using data from PEARL. The rationale behind this new development was the problems we had encountered with the classical measures of RA when applied to early arthritis patients. For its validation, we analyzed its feasibility, validity, and sensitivity to change.

Following the recommendations of OMERACT (Outcome Measures in Rheumatology) for core set outcomes measures in RA (Tugwell and Boers 1993; Tugwell and Boers 1993), we based our new index on the measures already included in other widely used measures: TJC, SJC, PGA and the acute phase reactants (ESR or CRP). We developed 10 alternative versions for the HUPI taking into account the distribution of each single item in our cohort (Table 2). To develop these 10 versions each single variable included was divided into quartiles, each of which was assigned an ordinal value from 0 to 3. Additionally, we defined different cutoff for TJC and ESR stratified by gender and different cutoff for PGA depending on age (older or younger than 40 years). The CRP level was scored using 2 strategies: one according to quartile distribution (CRP1) and the other according to theoretical thresholds based on local reference ranges (CRP2). In addition to the HUPI versions including only ESR, CRP1 level, or CRP2 level, we described 4 different possibilities to input the APR: (1) APR1 approach was calculated using the average of the scores of the ESR and CRP1; (2) APR2 approach was calculated using the average of the scores of the ESR and CRP2; (3) APR3 was calculated with the scores of ESR or CRP1 or both depending on which one was available; and (4) APR4 was calculated when only one of them (ESR or CRP) was available or with the average of their scores when both were available.

To determine which of these versions was the most reliable and performed the best we analysed each aspect of the validation process, namely feasibility, reliability, construct validity, and responsiveness.

**Feasibility** includes domains such as completion time, difficulty, clarity, and acceptance by both patients and clinicians. Feasibility was quantified by creating an ad hoc measure ranging from 0 (unfeasible) to 3 (completely feasible) to evaluate 3 domains: completion time—depending on the number of variables included—; clarity of the calculation—depending on the variables simplicity—; and acceptance—low probability of missing data—. Each investigator rated each index in the three domains independently. The final rating of the versions was the mean of the three ratings.

**Reliability** embraces the concepts of internal consistency and reproducibility. The internal consistency or “good construction” of each HUPI version was tested using Cronbach’s alpha (where  $\alpha \leq 0.70$  indicates that individual items provide an inadequate contribution to the overall scale, and values of  $\alpha \geq 0.90$  suggest redundancy).

**Construct validity** refers to the proximity of our measure to similar measures (convergent validity) and distance from dissimilar measures (divergent validity). When comparing the HUPI with similar construct measures (disease activity measures), a high correlation would be expected; when comparing it with less closely related constructs, such as function; a lower correlation would be expected.

**Criterion validity** was evaluated using receiver operating characteristic (ROC) curves with MDA (Wells, Boers et al. 2005; Wolfe, Rasker et al. 2007) as the external criterion. As previously explained, MDA was developed in 2005 by Wells et al as a satisfactory state of disease activity to compare different treatment a strategy, bearing in mind that true remission is difficult to achieve in patients with RA. Two equivalent definitions were formulated, one based on the DAS28 (EULAR response criteria) and the other based on meeting cut-offs in 5 of the 7 World Health Organization/International League of Associations for Rheumatology core set outcome measures, which is the set used in our analysis.

**Responsiveness**, also called sensitivity to change, is defined as “the ability of an instrument to accurately detect change when it has occurred” (de Bruin, Diederiks et al. 1997), implying that the intervention administered to the study patients involved an effect with a known direction. To study responsiveness we used different anchors of change as explained in the statistical methods below.

**Table 2:** Description of the 10 versions of the HUPI index.

HUPI-INDEX	Variables
HUPI-1	SJC + TJC + PGA1 + ESR
HUPI-2	SJC + TJC + PGA1 + CRP1
HUPI-3	SJC + TJC + PGA1 + APR1
HUPI-4	SJC + TJC + PGA2 + ESR
HUPI-5	SJC + TJC + PGA2 + CRP1
HUPI-6	SJC + TJC + PGA2 + CRP2
HUPI-7	SJC + TJC + PGA2 + APR1*
HUPI-8	SJC + TJC + PGA2 + APR2*
HUPI-9	SJC + TJC + PGA2 + APR3*
HUPI-10	SJC + TJC + PGA2 + APR4*

**Abbreviations:** SJC, swollen joint count; TJC, tender joint count; PGA1, global disease assessment by patient by age; PGA2, global disease assessment by patient irrespective of age; CRP1, C-reactive protein by quartile distribution; CRP2, C-reactive protein by theoretical thresholds based on local reference ranges; ESR, erythrocyte sedimentation rate.

\*APR1=(ESR+CRP1)/2; APR2=(ESR+CRP2)/2; APR3=(ESR ± CRP1)/2; APR4=ESR or CRP or (ESR+CRP2/2).

## Statistical procedures

All analyses were performed using Stata 10 for Windows (StataCorp LP, College Station, TX, USA). Baseline characteristics for patients included in each sub-analysis were summarised by median (IQR), mean (SD), or absolute and relative (%) frequencies, as appropriate.

To compare DAS28-ESR with DAS28-CRP, we generated a variable, "DIFDAS", as the subtraction of DAS28-CRP to DAS28-ESR. Pearson correlation tests were applied to determine whether there was any association between DIFDAS and independent categorical or continuous factors.

A multivariate linear regressions analysis was performed including all variables that were statistical significant at the bivariate analyses. Best fit models were obtained by stepwise backward estimation, removing all variables with a  $p > 0.05$ .

ROC analysis was performed to generate a new set of cut-off points for DAS28-CRP. Each cut-off point was selected based on the best trade-off values between sensitivity (Se) and specificity (Sp). ROC curves were also obtained to estimate whether the differences in the area under the curve between the indices were statistically significant. The same procedure was applied to obtain the cut-off value to consider MDA for both DAS28 indices.

To explore potential gender bias, we used Mann Whitney U test for continuous independent variables and chi-square test for qualitative variables. A  $p < 0.05$  was considered statistically significant in this sub-analysis.

Different statistics were applied to test each aspect of the validity of the new developed HUPI. Reliability was tested using Cronbach's alpha. Construct validity was evaluated through Pearson correlations. We tested the HUPI against the DAS28 and the SDAI, and then against the HAQ. To evaluate criterion validity the statistic applied was the area under the curve (AUC) (Hanley and McNeil 1982). The ROC curve of the HUPI was compared with that of the DAS28-ESR, the DAS28-CRP, and the SDAI. To evaluate responsiveness, the intervention was the treatment initiated by the physician, which in most instances was methotrexate. We analysed the AUC of the change in the HUPI for identifying patients who improved after 6 months of treatment.

Responsiveness was tested against three definitions of improvement as follows: 1) a change in the PhGA  $> 10$  between baseline and 6 months of follow-up; 2) the same definition but for PGA; and 3) change in the DAS28 compared with the change in the HUPI. Responsiveness by the first two definitions was tested comparing AUC of the ROC curves to determine statistically significant differences between the various indices. Responsiveness by the third definition was tested with the beta coefficient from the linear regression analysis. For this analysis statistical significance was set at a  $P$  value of less than 0.05; if Bonferroni correction was needed because of multiple comparisons, then the  $P$  value was set at less than 0.0125.

## **PART II – RESULTS**



## Chapter 1: Scopes and limitations of disease activity measures

As previously mentioned, DAS28 is the most widely used composite measure to evaluate RA patients. It is a simplified version of DAS to avoid more extensive and time-consuming joint counts. The DAS28 certainly improves the feasibility of the DAS without losing any significant information (Prevoo, van 't Hof et al. 1995). Both composite indices are calculated using a very complex formula and include ESR as acute phase reactant. ESR is extensively used to assess disease activity in RA but it can be influenced by several conditions (Gabay and Kushner 1999). Due to concerns regarding the specificity of ESR a CRP based version of the DAS28 was proposed (Fransen J 2004). The use of CRP levels has different advantages: laboratory tests are faster, it is more sensitive to short-term changes in disease activity (Kushner 1991) and its measurement can be standardized in a central laboratory for multicentre clinical trials. Alternatively, CRP is more accurate as indicator of inflammation (Deodhar 1989; Thompson, Milford-Ward et al. 1992; Gabay and Kushner 1999). Although it was originally believed that there was a very good correlation between the DAS28-CRP and the DAS28-ESR, some authors argue that the DAS28-CRP may need lower cut-offs for categorizing disease activity (Inoue, Yamanaka et al. 2007). For this reason we designed an initial evaluation to determine which factors might account for the differences between the two versions of the DAS28 and to what extent these factors had any consequences into the assessment of RA activity in daily clinical practice.

DAS28-CRP has not been formally validated and the accepted cut-off points established for DAS28-ESR do not seem to perform well for DAS28-CRP (Inoue, Yamanaka et al. 2007). After showing that both indexes are not fully equivalent, ought to the DAS28-ESR tendency to yield higher values, especially in women and patients with long standing disease, we were interested in generating specific cut-off points for DAS28-CRP based on patient evaluations.

Another important aspect that can arise when using composite index is the possible gender bias. For this reason, we were interested in exploring the influence of gender on treatment response assessed by DAS28 as a composite index, and by the doctor and the patient's assessment in our cohort.

The following articles are presented:

1. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? (Clin Exp Rheumatol 2008;26:769-75)
2. Estimated cut-off points for the 28-joint disease activity score based on C-reactive protein in a longitudinal register of early arthritis. (J Rheumatol 2010;37:1439-43).
3. Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid. (Reumatol Clin 2010;6:134-40)

**1.1. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? (Clin Exp Rheumatol 2008;26:769-75)**

# Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis?

I. Castrejón<sup>1</sup>, A.M. Ortiz<sup>1</sup>, R. García-Vicuña<sup>1</sup>, J.P. Lopez-Bote<sup>1</sup>, A. Humbría<sup>1</sup>,  
L. Carmona<sup>2</sup>, I. Gonzalez-Alvaro<sup>1</sup>

<sup>1</sup>Rheumatology Service, Hospital Universitario de la Princesa, Madrid;

<sup>2</sup>Research Unit, Fundación Española de Reumatología, Madrid, Spain.

---

## Abstract

*A formula for calculating disease activity score with 28 joint counts (DAS28) with C-reactive protein (CRP) instead of the erythrocyte sedimentation rate (ESR) has been proposed.*

---

## Objective

*Here we analyze the factors that contribute to the differences in the DAS28 when calculated using either the ESR (DAS28-ESR) or the CRP values (DAS28-CRP).*

---

## Methods

*We analyzed the data from 587 visits made by 220 patients with early arthritis. The age at the onset of the disease was 51±16 years old and 76.3% of the patients were women. The disease evolution at the first visit was 5 months and at each visit information related to several variables was collected, including that necessary to calculate the DAS28-ESR and DAS28-CRP. We defined a new variable DIFDAS=DAS28-ESR – DAS28-CRP to analyze which independent variables account for differences between the two indexes.*

---

## Results

*There was a correlation between the two indexes of 0.91 ( $p<0.0001$ ), although the DAS28-ESR value obtained was higher than that of DAS28-CRP at approximately 90% of the visits. Significantly, the difference between both indexes was higher than 0.6 in 44% of the visits studied. A multivariate analysis showed that female gender and disease duration were associated with the higher values obtained for DAS28-ESR when compared to those of DAS28-CRP.*

---

## Conclusion

*Our data show that DAS28-ESR and DAS28-CRP are not fully equivalent, because the former usually produces higher values. This finding is particularly relevant in females and patients with a long disease duration.*

---

## Key words

Rheumatoid arthritis, erythrocyte sedimentation rate, C-reactive protein, outcome measures.

Isabel Castrejón, MD  
 Ana M. Ortiz, MD, PhD  
 Rosario García-Vicuña, MD, PhD  
 Juan P. Lopez-Bote, MD, PhD  
 Alicia Humbría, MD, PhD  
 Loreto Carmona, MD, PhD  
 Isidoro Gonzalez-Alvaro, MD, PhD

This study was funded by grants from the Instituto de Salud Carlos III: FIS 05/2041 and 05/2044 to IGA.

Please address correspondence and reprint requests to:

Dr. Isidoro Gonzalez-Alvaro,  
 Rheumatology Service,  
 Hospital Universitario de la Princesa,  
 c/ Diego de León 62, 28006 Madrid,  
 Spain.

E-mail: isidoro.ga@ser.es

Received on August 20, 2007; accepted in revised form on January 22, 2008.

© Copyright CLINICAL AND

EXPERIMENTAL RHEUMATOLOGY 2008.

## Introduction

Rheumatoid arthritis (RA) may cause severe and irreversible joint destruction leading to functional disability, impaired quality of life and increased comorbidity and mortality (1). Lately, the development of new therapeutic strategies has improved the prognosis of patients with RA. However such therapies are not free of serious adverse events and they are also very costly, posing a substantial economic burden to health care systems (2). It is therefore necessary to establish an accurate risk/benefit ratio for these agents as well as to quantify patients' response to these treatments, both in clinical trials and in daily practice.

The DAS is a combined index that incorporates, in a continuous score, the Ritchie articular index, the 44 swollen joints count, the global disease activity as assessed by the patient, and the erythrocyte sedimentation rate (ESR) as an acute phase reactant. The DAS has been proved to be a sensitive and specific tool to measure disease activity in RA (3). Moreover, when DAS is used to monitor treatment response together with monthly visits, the level of disease control achieved is greater than when more traditional schedules are utilized (4).

One of the major downsides of the DAS is that its joint counts are time consuming. A simplified version of the DAS that uses 28-joint counts, the DAS28, was then developed. The DAS28 certainly improves the feasibility of the index without losing any significant information of the original DAS (5, 6). An additional problem associated with the DAS or DAS28 index is that although the ESR is extensively used to assess disease activity in RA, it can be influenced by several conditions such as age, female gender, anemia, serum fibrinogen levels, immunoglobulins and rheumatoid factor (7). C-reactive protein (CRP) is more accurate as indicator of inflammation than ESR and it is also more sensitive to short-term changes (7-9). Discrepancies between ESR and CRP values may result from the effect of blood constituents that are not related to inflammation but that can interfere with the ESR. Accordingly, a

formula for DAS28 has been proposed whereby the index is calculated using CRP instead of ESR (<http://www.das-score.nl>). Although it was originally believed that there was a very good correlation between the DAS28-CRP and the DAS28-ESR, some authors argue that the DAS28-CRP may need lower cut-offs for categorizing disease activity (10, 11).

In the view of the above, the aim of our study was to determine which factors might account for the differences between the two versions of the DAS28 and to what extent these factors had any consequences into the assessment of RA activity in daily clinical practice.

## Patients and methods

This is a prospective longitudinal observational study in which all patients attending the Early Arthritis Clinic in our center from September 2001 to June 2006 were included. To be referred to the clinic, patients had to have two or more swollen joints for at least four weeks and symptoms for less than one year. Patients were excluded if they had been diagnosed of gouty arthritis, septic arthritis, spondyloarthropathies or connective tissue diseases during the follow-up. The study protocol was reviewed and approved by the Local Research Ethics Committee and all patients who entered the study signed a written consent form after being informed about the details of the protocol.

At each visit, the following data are collected per protocol and entered into an electronic database: clinical and demographic information, data about treatments with disease modifying anti-rheumatic drugs, 28 tender and swollen joint counts (TJC and SJC, respectively), global disease activity on a 100 mm visual analogue scale assessed both by the patient (GDAP) and by the physician (GDAPh), and basic laboratory tests including ESR and CRP. DAS28 indexes, with ESR and with CRP were calculated as previously described:

$$\text{DAS28-ESR} = 0.56 \cdot \sqrt{\text{TJC28}} + 0.28 \cdot \sqrt{\text{SJC28}} + 0.70 \cdot \ln(\text{ESR}) + 0.014 \cdot (\text{GDAP}) \quad (5)$$

$$\text{DAS28-CRP} = 0.56 \cdot \sqrt{\text{TJC28}} + 0.28 \cdot \sqrt{\text{SJC28}} + 0.36 \cdot \ln(\text{CRP}+1) + 0.014 \cdot (\text{GDAP}) + 0.96$$

(<http://www.das-score.nl>)

## Conflict of interest:

Dr. I. Gonzalez-Alvaro has received unrestricted research funding from Abbott Laboratories, Sanofi-Aventis and Bristol-Myers Squibb. However, these research projects have no relation with this study; the other co-authors have declared no competing interests.

### Statistical procedures

All statistical analyses were performed using Stata 9.2 for Windows (StataCorp LP, College Station, TX, USA). We first compared the distribution of DAS28-ESR and DAS28-CRP using graphic tools such as the *kdensity* command that provides kernel density estimations. Then, to compare how both indexes evaluated disease activity at each visit, we created the DIFDAS variable:

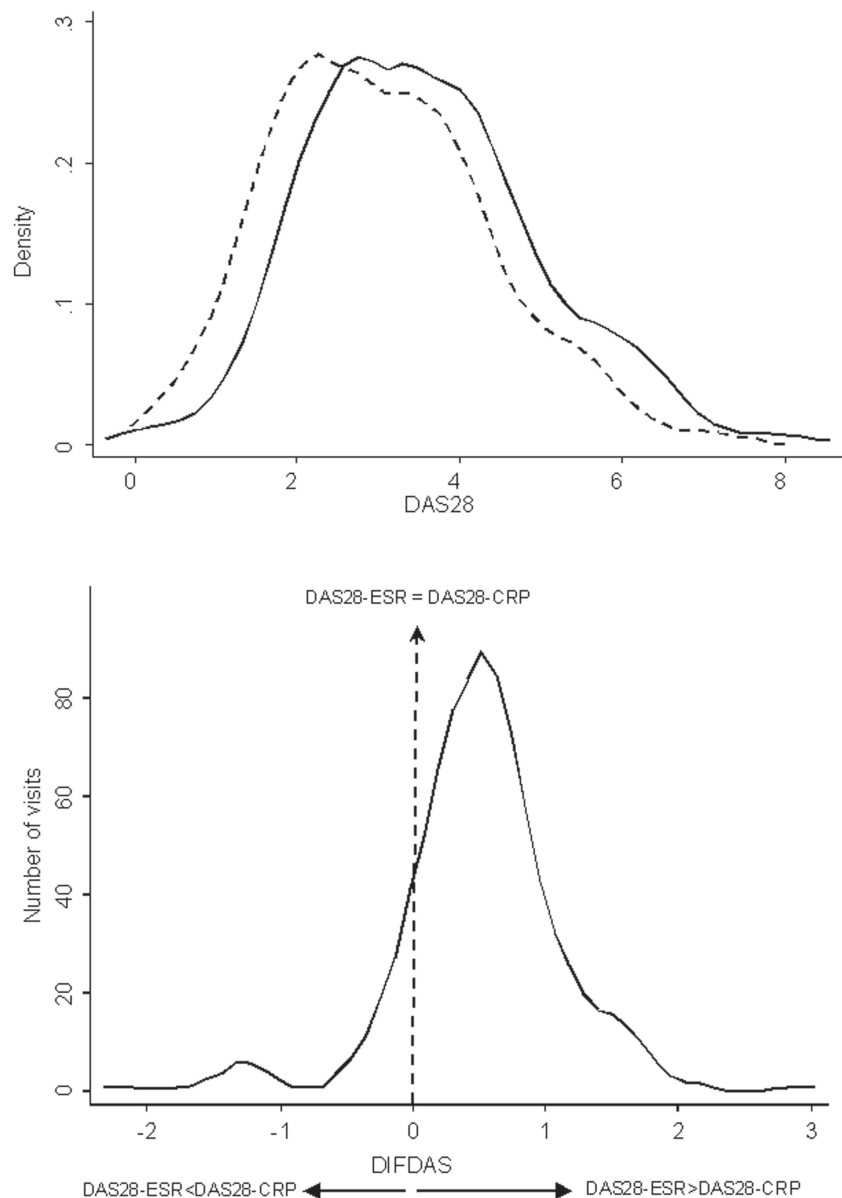
$$\text{DIFDAS} = \text{DAS28-ESR} - \text{DAS28-CRP}$$

We used the Mann-Whitney-U or Pearson correlation tests to determine whether there was any association between DIFDAS and independent categorical or continuous factors, respectively. Then, we undertook multivariate linear regressions by using the *glm* command of Stata (Gaussian as *family* option and identity as *link* option) including all variables that reached a  $p < 0.05$  at the bivariate analyses. Best fit models were obtained by stepwise backward estimation, removing all variables with a  $p > 0.05$ .

### Results

A total of 220 patients (76.4% female) were included in the study. We analyzed the data from 587 structured visits in a follow-up period of two years, including 220 initial visits, 139 second visits, 125 third visits and 103 fourth visits after 6, 12 and 24 months of follow-up, respectively. The age at the onset of the disease was  $51 \pm 16$  years old and the mean disease duration at the first visit  $5.1 \pm 2.9$  months. Rheumatoid factor was positive in 128 patients (58.2%) and anti-cyclic citrullinated peptide antibodies in 70 (31.2%) of them. They represent a fairly average early arthritis cohort, with more than half of the patients (57.3%) already fulfilling ACR criteria for RA at entry.

The values of DAS28-ESR in our cohort ranged from 0 to 8.2, with a median of 3.4 and an IQR of 2.5–4.4. In contrast, the DAS28-CRP values ranged from 0.2 to 7.7 and showed a median of 2.8 (IQR: 1.9–3.9). Although the correlation coefficient of the two indexes was 0.91 ( $p < 0.001$ ), the value of DAS28-ESR was higher than DAS28-CRP in approximately 90% of the visits. The distribution of the values for both indexes



**Fig. 1.** A: Distribution of the DAS28 index calculated with the ESR (solid line) and CRP (dotted line) values in our population. B: Distribution of the DIFDAS values (see *Materials and methods* for definition) at all the visits analyzed in our cohort.

produced similar shapes, although the distribution curve for the DAS28-ESR was displaced toward higher values (Fig. 1A).

The difference between both indexes was higher than 0.6 at 44% of the visits and higher than 1.2 at 26% of the visits. These represent the minimum relevant variation in response to treatment depending on initial DAS28 value (less than 5.1 or higher than 5.1 respectively) according to the EULAR criteria (12). Furthermore, considering the cut-off points proposed by Prevoo *et al.* (5) our patients were in remission at 41%

of the visits when the DAS28-CRP was applied but only in 26% of the visits when applying the DAS28-ESR (Table I). Conversely, the proportion of cases with low, moderate or high disease activity was higher when the DAS28-ESR was applied than when the DAS28-CRP (Table I).

Then, we studied which variables could explain these differences by generating a new variable, DIFDAS, as defined in the Methods. The distribution of DIFDAS was displaced towards positive values, indicating that DAS28-ESR renders a higher value than DAS28-CRP in most

**Table I.** Classification of disease activity.

Disease activity	Remission	Low	Moderate	High
DAS28-CRP	41.3%	14.4%	35%	9.3%
DAS28-ESR	26%	17.9%	40.8%	15.2%

visits (Fig. 1B). As shown in Table II, several variables were associated with differences in DIFDAS in the bivariate analysis. However, after adjusting for the ESR and CRP values, the multivariate analysis demonstrated that the variables of gender and disease duration were those that contributed significantly to the differences between DAS28-ESR and DAS28-CRP. Indeed, women showed higher DAS28-ESR values than those obtained with the DAS28-CRP (Fig. 2), and the contribution of gender to the differences between both indexes was 0.2 points higher for women in the DAS28-ESR index (Table II).

With regards to disease duration, the multivariate analysis suggested that the difference in the DAS28-ESR increased with respect to the DAS28-CRP, with a regression coefficient of 0.02 per month of disease duration (Table II), although this effect was not linear (Fig. 3A). The differences between the two indexes was 0.5 points on average in the first 20

months of disease duration and thereafter, the differences increased continuously over time due to the increasing relative DAS28-ESR values. Regarding the clinical consequences of this finding, it seems that after two years of disease duration, disease activity estimated by DAS28-ESR tends to reach a plateau (Fig. 3B) whereas when estimated with CRP the disease activity of the cohort continues to improve (Fig. 3C).

Nevertheless, since the DAS score was developed and validated to be used in rheumatoid arthritis patients, we reanalyzed our data separately both in the RA patients and in patients with undifferentiated arthritis (UA). Coefficients for gender were  $0.19 \pm 0.06$  ( $p=0.002$ ) for RA patients and  $0.26 \pm 0.09$  ( $p=0.005$ ) for UA patients and for disease duration were  $0.018 \pm 0.002$  ( $p<0.001$ ) for RA and  $0.021 \pm 0.002$  ( $p<0.001$ ) for UA patients. In both cases, the coefficients were like those described for the whole population (Table II).

## Discussion

A formula to calculate DAS28 using the CRP values as the acute phase reactant variable has been proposed on the basis that CRP response to treatment is faster than ESR response (13, 14). Accordingly, DAS28-CRP has been included as an outcome variable in some clinical trials, although there is still insufficient information about how the DAS28-CRP index behaves in comparison with DAS28-ESR. In this regard, this study shows that the DAS28-ESR and DAS28-CRP values are not interchangeable and that DAS28-ESR tends to produce higher values in women and long-term disease patients.

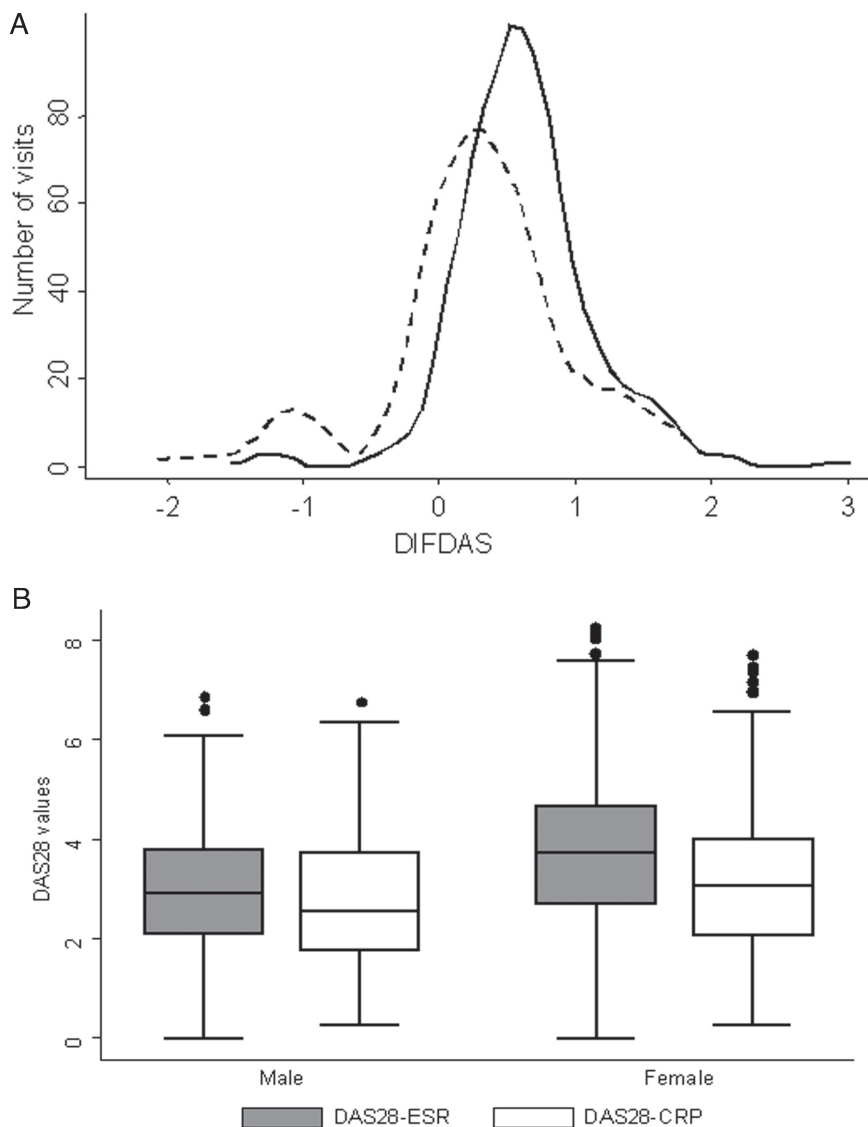
With regards the first of these issues, in 44% and 26% of the visits analyzed in our study the differences between both indexes were greater than 0.6 and 1.2 points respectively, the minimal improvements considered to be relevant by EULAR criteria depending on baseline DAS28 measurement (12). Therefore, our data suggest that despite the good correlation between both indexes, if we evaluate disease activity with the DAS28-ESR in one visit and then with DAS28-CRP in the next, we may incorrectly consider the patient to

**Table II.** Bivariate and multivariate analysis.

	Bivariate analysis		Multivariate analysis	
	DIFDAS	p-value	Coef. of regression	p-value
Gender (male vs. female)	0.24 [-0.10-0.52] / 0.58 [0.35-0.86]	<0.05	0.199	0.001
Disease (RA vs. UA)	0.51 [0.20-0.83] / 0.52 [0.23-0.80]	0.82		
Therapy: None	0.49 [0.17-0.75]		0.110	0.011
MT	0.51 [0.24-0.80]	0.015		
CT	0.63 [0.34-1.17]		0.142	0.017
CCP (+) vs. (-)	0.49 [0.17-0.77] / 0.53 [0.28-0.94]	0.006	-----	-----
RF (+) vs. (-)	0.54 [0.26-0.96] / 0.53 [0.19-0.96]	0.49	-----	-----
GDAP	0.0092	0.828	-----	-----
GDAPh	-0.0005	0.991	-----	-----
Pain	0.034	0.415	-----	-----
HAQ	0.118	0.005	-----	-----
ESR	0.381	<0.001	0.23	0.001
CRP	-0.118	0.005	-0.84	0.001
Age (years old)	0.007	0.862	-----	-----
Disease duration	0.402	<0.001	0.024	0.001

RA: rheumatoid arthritis; UA: undifferentiated arthritis; MT: monotherapy; CT: combined therapy; CCP: anti-cyclic citrullinated peptide antibodies; RF: rheumatoid factor; GDAP: global disease activity assessment by the patient; GDAPh: global disease activity assessment by the physician; HAQ: health assessment questionnaire; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.





**Fig. 2.** Effect of gender on the evaluation of disease activity estimated with ESR and CRP. **A:** Distribution of the DIFDAS variable in female (solid line) or male patients (dotted line: see *Materials and methods* for definition). **B:** Distribution of the DAS28 values calculated with ESR (grey boxes) or CRP (white boxes) according to gender. Data are presented as the interquartile range (p75 upper edge of the box, p25 lower edge, p50 midline in the box), as well as the p95 (upper line from the box) and p5. Dots represent the outliers.

have improved. In addition, if we apply the cut-off values proposed for the DAS28-ESR to DAS28-CRP, we might underestimate the disease activity of the patients and increase the proportion of patients in remission or with weak disease activity. Indeed, the estimated cut-off values proposed by Inoue *et al.* for DAS28-CRP (10) support our finding that the DAS28-CRP values are, on average, 0.5 points lower than the DAS28-ESR values.

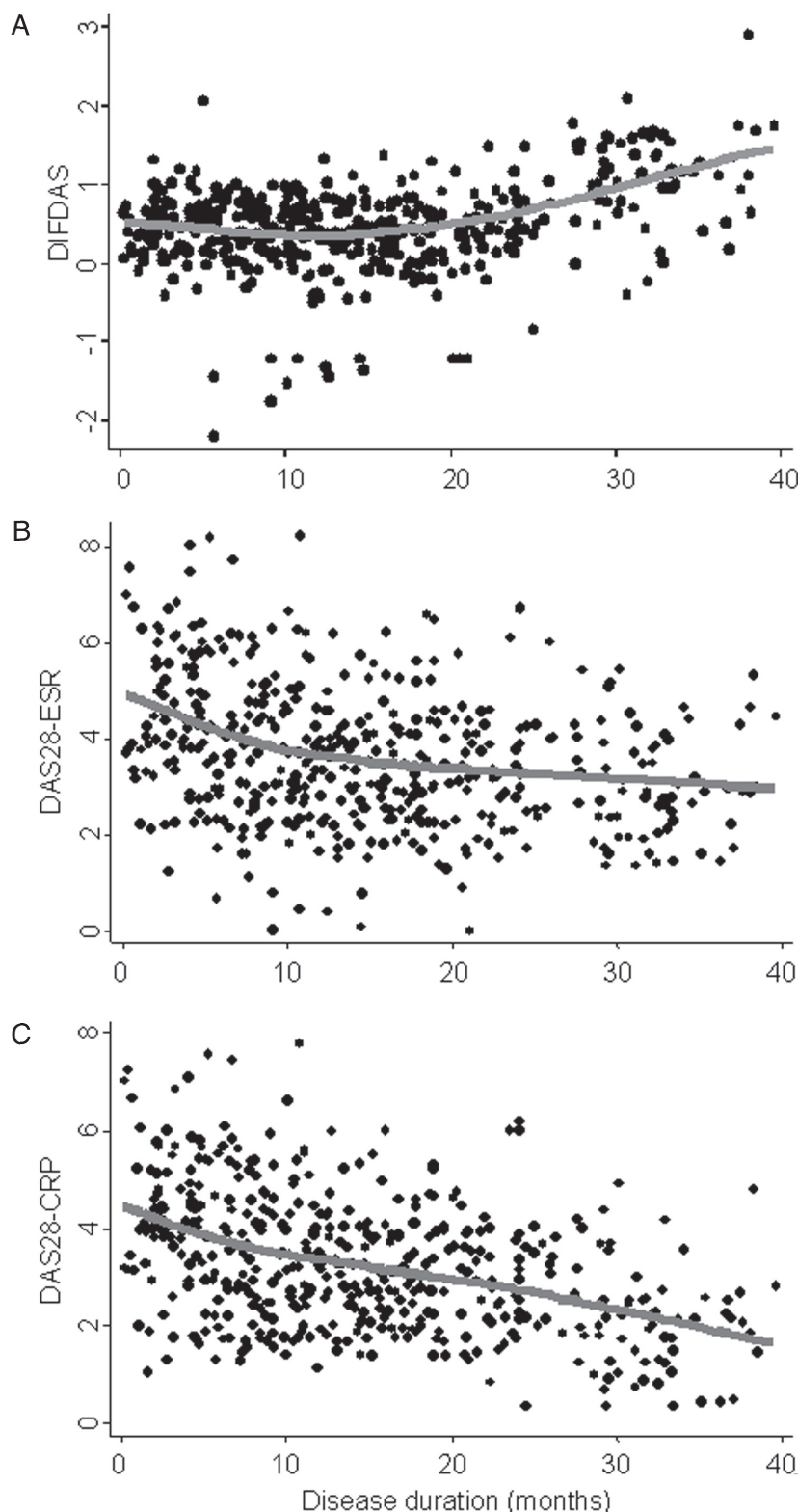
The second finding raises the question as to which index is best, as both have

their advantages and disadvantages. For clinical trials, we would expect to use an index that rapidly shows the effect of therapeutic agents. Considering the fast response of CRP to variations in disease activity, we may choose DAS28-CRP in this case. On the other hand, in daily clinical practice we would prefer an index that showed us how the patient was on average during the preceding period, and DAS28-ESR would probably be better in this respect due to its slower response to variations in disease activity when compared with

CRP. However, DAS28-ESR has additional problems as it is less sensitive to changes in long-term patients. In addition, women may be less frequently considered in remission when assessing disease activity with the DAS28-ESR ( $\text{DAS28} < 2.6$ ).

On the other hand, regarding factors that might bias our results, we may consider that only about 60% of our patients fulfilled the ACR criteria for RA classification (15). However, this factor did not significantly account for the differences between DAS28 calculated with ESR or CRP. In addition, CRP levels may increase with age in men but not in women (16). Nevertheless, our results suggest that this effect is probably moderate and it is clearly less important than the enhanced ESR levels observed in females with a long-term disease. Furthermore, other factors such as race, smoking, increased blood pressure, diabetes, high body mass index or abdominal adiposity may also be associated with increased CRP levels (16, 17). In our analysis, we did not adjust for all these variables since this information was not collected, and therefore, we can not exclude that they may influence the final results. Thus, perhaps the indexes proposed to evaluate the disease activity in RA patients should be evaluated in different subsets of patients in order to establish how robust they are.

In summary, our data suggest that when the DAS28 is calculated with the CRP it may be more accurate to determine RA activity, especially in long-term female patients. However, specific cut-off points should be estimated for the DAS28-CRP since it produces lower values than DAS28-ESR (10, 11). In this regard, preliminary threshold values have been proposed for DAS28-CRP in Japanese patients (10) and our group is now involved in a study to estimate such cut-off points in our population. In addition, our data suggest that DAS28 might behave similarly when applied to RA or UA patients. In this regard, we believe that comparisons between different populations should provide additional information about the reliability and reproducibility of these indexes.



**Fig. 3.** Effect of disease duration on the evaluation of disease activity estimated with ESR or CPR. **A:** Distribution of the DIFDAS values (see *Materials and methods* for definition) over the disease duration of the patients in the cohort. **B:** Evolution of disease activity of the cohort estimated with DAS28-ESR. **C:** Evolution of disease activity of the cohort estimated with DAS28-CRP. Dots represent the values of each variable for each patient in the cohort at the different follow-up visits. The solid grey line represents the local (bandwidth 0.8) least-squares weighted regression of each variable (DIFDAS, DAS28-ESR or DAS28-CRP) with respect to disease duration estimated using the command `lowses` of Stata 9.2 for Windows.

## Acknowledgements

The authors wish to recognize the influence of Dr Armando Laffon on our work, who instilled us with the eagerness to learn from the patient, both in the consultancy and in the laboratory, so that our findings might further improve the medical attention that we can offer them.

## References

1. PINCUST T, CALLAHAN LF, SALE WG, BROOKS AL, PAYNE LE, VAUGHN WK: Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984; 27: 864-72.
2. DOAN QV, CHIOU CF, DUBOIS RW: Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis. *J Manag Care Pharm* 2006; 12: 555-69.
3. VAN DER HEIJDE DM, VAN'T HOF MA, VAN RIEL PL *et al.*: Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49: 916-20.
4. GRIGOR C, CAPELL H, STIRLING A, MCMAHON AD, LOCK P, VALLANCE R *et al.*: Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004; 364: 263-9.
5. PREVOO ML, VAN'T HOF MA, KUPER HH, VAN LEEUWEN MA, VAN DE PUTTE LB, VAN RIEL PL: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 44-8.
6. FRANSEN J, VAN RIEL PL: DAS remission cut points. *Clin Exp Rheumatol* 2006; 24: S29-32.
7. GABAY C, KUSHNER I: Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med* 1999; 340: 448-54.
8. DEODHAR SD: C-reactive protein: the best laboratory indicator available for monitoring disease activity. *Cleve Clin J Med* 1989; 56: 126-30.
9. THOMPSON D, MILFORD-WARD A, WHICHER JT: The value of acute phase protein measurements in clinical practice. *Ann Clin Biochem* 1992; 29: 123-31.
10. INOUE E, YAMANAKA H, HARA M, TOMATSU T, KAMATANI N: Comparison of Disease Activity Score (DAS)28- erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values. *Ann Rheum Dis* 2007; 66: 407-9.
11. MATSUIT T, KUGA Y, KANEKO A *et al.*: Disease Activity Score 28 (DAS28) using C-reactive protein underestimates disease activity and overestimates EULAR response criteria compared with DAS28 using erythrocyte sedimentation rate in a large observational cohort of rheumatoid arthritis patients in Japan. *Ann Rheum Dis* 2007; 66: 1221-6.



12. VAN GESTEL AM, HAAGSMA CJ, VAN RIEL PL: Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998; 41: 1845-50.
13. WALSH L, DAVIES P, MCCONKEY B: Relationship between erythrocyte sedimentation rate and serum C-reactive protein in rheumatoid arthritis. *Ann Rheum Dis* 1979; 38: 362-3.
14. WARD MM: Relative sensitivity to change of the erythrocyte sedimentation rate and serum C-reactive protein concentration in rheumatoid arthritis. *J Rheumatol* 2004; 31: 884-95.
15. ARNETT FC, EDWORTHY SM, BLOCH DA *et al.*: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-24.
16. KRITCHEVSKY SB, CESARI M, PAHOR M: Inflammatory markers and cardiovascular health in older adults. *Cardiovasc Res* 2005; 66: 265-75.
17. KHERA A, MCGUIRE DK, MURPHY SA *et al.*: Race and gender differences in C-reactive protein levels. *J Am Coll Cardiol* 2005; 46: 464-9.

**1.2. Estimated cut-off points for the 28-joint disease activity score based on C-reactive protein in a longitudinal register of early arthritis. (J Rheumatol 2010;37:1439-43).**

# Estimated Cutoff Points for the 28-Joint Disease Activity Score Based on C-reactive Protein in a Longitudinal Register of Early Arthritis

ISABEL CASTREJÓN, ANA M. ORTIZ, ESTHER TOLEDANO, SANTOS CASTAÑEDA, ALBERTO GARCÍA-VADILLO, ESTHER PATIÑO, and ISIDORO GONZÁLEZ-ÁLVARO

**ABSTRACT. Objective.** To estimate the cutoff points for the 28-joint Disease Activity Score (DAS28) calculated using C-reactive protein (CRP) measurements from patients with early arthritis.

**Methods.** We analyzed data from 568 visits of 207 patients enrolled in our prospective longitudinal register of early arthritis. Six rheumatologists evaluated the degree of disease activity at each visit on the basis of the available clinical data, and the final degree of disease activity was established by consensus. DAS28 values were calculated for each visit using CRP or erythrocyte sedimentation rate (ESR). Through a ROC analysis, cutoff points for both indices, as well as for minimal disease activity (MDA), were selected on the basis of the best tradeoff values between sensitivity and specificity.

**Results.** The cutoff values to classify disease activity with the DAS28-CRP were 2.3, 3.8, and 4.9, considering remission at < 2.3, low disease activity 2.3–3.8, moderate disease activity 3.8–4.9, and high disease activity > 4.9. The cutoff value for MDA when calculated with CRP was 2.6. The area under the ROC curves was always greater for DAS28-CRP than for DAS28-ESR, reaching statistical significance for low/moderate activity and for the MDA.

**Conclusion.** Our study confirms that the cutoff points for DAS28-CRP are lower than those described for DAS28-ESR, suggesting that DAS28-CRP may be more accurate to assess disease activity in our population. (J Rheumatol First Release May 15 2010; doi:10.3899/jrheum.091333)

## Key Indexing Terms:

DISEASE ACTIVITY SCORE 28  
UNDIFFERENTIATED ARTHRITIS

RHEUMATOID ARTHRITIS  
DISEASE ACTIVITY

OUTCOME  
REMISSION

The management of rheumatoid arthritis (RA) has improved greatly in the last 2 decades, and the availability of new drugs, particularly the biological agents, has helped achieve better control of this disorder. However, the development of other clinical tools such as the Disease Activity Score (DAS)<sup>1</sup> may also help to improve the control of RA<sup>2</sup>. Indeed, close followup of treatments in accord with predetermined DAS values may be decisive to improving radiographic outcome, physical function, and quality of life, compared with traditional management strategies<sup>3</sup>.

From the Rheumatology Department, Hospital Universitario la Princesa, Madrid, Spain.

Supported by the Instituto de Salud Carlos III (grant numbers FIS 05/2044 and 05/2041 to Dr. García-Vadillo). Dr. González-Álvarez's work is funded in part by a grant from the Instituto de Salud Carlos III (Intensificación de la Labor de Investigación) to promote research in the Spanish National Health Service. Supported in part by RETICS Program, RD08/0075 (RIER) from Instituto de Salud Carlos III (ISCIII).

E. Toledano, MD; A.M. Ortiz, MD, PhD; I. Castrejón, MD, current address Rheumatology Department, Hospital Clínico San Carlos, c/ Profesor Martín Lagos s/n, 28040 Madrid, Spain; S. Castañeda, MD, PhD; A. García-Vadillo, MD, PhD; E. Patiño, MD, PhD; I. González-Álvarez, MD, PhD, Rheumatology Department, Hospital Universitario la Princesa.

Address correspondence to Dr. I. González-Álvarez, Rheumatology Department, Hospital Universitario la Princesa, c/ Diego de León 62, 28006 Madrid, Spain. E-mail: isidoro.ga@ser.es

Accepted for publication February 23, 2010.

While the DAS is a useful tool in clinical trials and observational studies, applying it in daily clinical practice is complicated due to the time required to perform joint counts. As such, there is no clear evidence regarding its efficacy in daily practice. Accordingly, a simplified version of the DAS has been described that uses 28-joint counts and is more feasible to implement, without losing significant information regarding the original score<sup>4</sup>. Due to concerns regarding the specificity of the erythrocyte sedimentation rate (ESR) as an acute-phase reactant, particularly as it may be influenced by unrelated factors such as age or gender, a C-reactive protein (CRP)-based version of the DAS28 was recently proposed<sup>5</sup>. The use of CRP levels has different advantages: laboratory tests are faster, CRP is more sensitive to short-term changes in disease activity<sup>6</sup>, and its measurement can be standardized in a central laboratory for multicenter clinical trials. The main concern is that DAS28 based on CRP (DAS28-CRP) has not been formally validated, and the accepted cutoff points established for DAS28 using the ESR (DAS28-ESR) do not seem to perform well for DAS28-CRP<sup>7</sup>. Indeed, we recently showed that both indexes are not fully equivalent since the DAS28-ESR tends to provide higher values than DAS28-CRP, especially in women and patients with long disease evolution<sup>8</sup>.

Therefore, we were interested in generating specific cutoff points for DAS28-CRP based on patient evaluations. We set out to define the cutoff points for DAS28-CRP in a population of patients from our early arthritis register.

## MATERIALS AND METHODS

We analyzed data from a prospective longitudinal observational study based on a register that includes all patients attending the early arthritis clinic (EAC) at our center. Our catchment area covers a population of 500,000 inhabitants, of whom more than 90% are covered by public health insurance. In addition, all primary care physicians in the area are aware of the EAC and how to refer patients to it.

To be referred to the clinic, patients must have had 2 or more swollen joints for at least 4 weeks and symptoms for less than a year. Patients diagnosed with gouty arthritis, septic or viral arthritis, osteoarthritis, spondyloarthropathies, or connective tissue diseases during the followup period were excluded from the study. Thus, we included only data from patients that fulfilled American College of Rheumatology (ACR) criteria for the diagnosis of RA<sup>9</sup> or patients with chronic undifferentiated arthritis. The protocol for the register included 4 visits during a followup period of 2 years and it was reviewed and approved by the local research ethics committee. Prior to entry into the register, all patients signed a written informed consent form.

Our study focused on the visits that took place between September 2001 and June 2006. At each visit, the following data were collected according to an established protocol and entered into an electronic database: clinical and demographic information; disease duration at the beginning of followup; data on treatment with disease modifying antirheumatic drugs (DMARD) and steroids; 28-joint count of tender (TJC) and swollen joints (SJC); global disease activity on a 100 mm visual analog scale assessed both by the patient (GDAP) and the physician (GDAPh); and the Spanish version of the Health Assessment Questionnaire (HAQ)<sup>10</sup> and basic laboratory tests including the evaluation of the ESR (Westergren method) and CRP (nephelometry). The DAS28 indices, both with ESR and CRP, were calculated as described (<http://www.das-score.nl>):

$$\begin{aligned}\text{DAS28-ESR} &= 0.56 * \sqrt{(\text{TJC28})} + 0.28 * \sqrt{(\text{SJC28})} + 0.70 * \ln(\text{ESR}) + \\ &\quad 0.014 * (\text{GDAP})^4 \\ \text{DAS28-CRP} &= 0.56 * \sqrt{(\text{TJC28})} + 0.28 * \sqrt{(\text{SJC28})} + 0.36 * \ln(\text{CRP} + 1) + \\ &\quad 0.014 * (\text{GDAP}) + 0.96\end{aligned}$$

Minimal disease activity (MDA) was considered, as described by Wells, *et al*<sup>11</sup>, when patients met 5 of the following criteria: pain (0–10) ≤ 2; SJC (0–28) ≤ 1; TJC (0–28) ≤ 1; HAQ (0–3) ≤ 0.5; GDAPh (0–10) ≤ 1.5; GDAP (0–10) ≤ 2; and ESR ≤ 20.

**Disease activity assessment.** Six rheumatologists were involved in assessing disease activity at each visit: 4 senior rheumatologists and 2 fellows from our department (3 of whom were female); the median experience in assessing RA patients was 13.2 years (range 2–24 yrs). The degree of disease activity was classified as remission, low activity, moderate activity, or high activity at each visit, based on TJC, SJC, GDAP, HAQ, ESR, and CRP. Because assessment was retrospective, the physicians could not physically examine the patients. In addition, they were blind to the GDAPh obtained at each visit. It is noteworthy that there was total agreement between these rheumatologists for 476 visits, while in the remaining 92 visits, the disease activity status was established by consensus of the evaluators. As such, the patients were considered to be in remission at 104 visits, while low, moderate, or high disease activities were considered in 289, 126, and 49 visits, respectively. We analyzed variables that influence disagreement at some visits using logistic regression.

**Statistical analysis.** We analyzed interobserver agreement of disease activity evaluation using the kappa command of Stata 9.2<sup>®</sup> for Windows (StataCorp LP, College Station, TX, USA). Two logistic regression models

were applied to assess independent variables related to the level of disagreement. The first model included level of disease activity assessed by the rheumatologists and the second model excluded this variable.

A receiver operating characteristic (ROC) analysis was performed on the data using the “roctab” command of Stata. Each cutoff point was selected on the basis of the best tradeoff values between sensitivity and specificity. ROC curves were also obtained with the roctab command of Stata, using the “graph” option. To estimate whether differences in the area under the curve (AUC) between indices were statistically significant, we used the Stata “roccomp” command that provides a test for the equality of the AUC using an algorithm described by DeLong, *et al*<sup>12</sup>. Statistical significance was accepted if the p value was less than 0.05.

The same procedure was followed to obtain the cutoff value to consider MDA for both DAS28 indices.

## RESULTS

**Patient characteristics.** A total of 568 visits by 207 patients enrolled in the register were analyzed in this study, 76.4% of whom were female. The mean age at the onset of the disease was 51 ± 16 years and the median disease duration at the first visit was 6 months (interquartile range 3.6–9). A more detailed description of this population has been published<sup>8</sup>.

**Agreement on evaluation of disease activity.** The kappa index of the information recorded by the 6 rheumatologists indicated that the best agreement between physicians was observed for the state of remission (κ = 0.65), followed by that of high activity (κ = 0.52). However, there was only moderate concordance in the intermediate degrees of disease activity (κ = 0.30–0.41).

These values reflect the variability between physicians in their perception of disease activity, as evident in Table 1 and as confirmed with a multivariate logistic regression model where disagreement was clearly more striking for the moderate level of disease activity (Table 2, model 1). Moreover, TJC and HAQ contributed to the disagreement irrespective of the moderate level of disease activity (Table 2, model 2). Intriguingly, the characteristics (gender, age, years of experience) of the evaluators did not contribute to the disagreement in the perception of disease activity, suggesting that highly individual variability exists.

**Estimation of the cutoff points for DAS28-CRP.** The best threshold values of the DAS28-CRP to stratify the patients in our population according to the state of their disease (in remission, low, moderate, or high disease activity) were 2.3 [sensitivity (Se) 87%; specificity (Sp) 96%], 3.8 (Se 78%; Sp 88%), and 4.9 (Se 84%; Sp 83%). These differed from the cutoff points obtained for DAS28-ESR of 2.7, 4.3, and 5.5, respectively. We analyzed the ROC curves for DAS28-ESR and DAS28-CRP in the patients considered in remission/low activity, low/moderate activity, and moderate/high activity to determine which score best evaluates disease activity. The areas under the ROC curves were always greater for DAS28-CRP than for DAS28-ESR at each level of disease activity, although statistically these differences were only significant for the low/moderate activity group (Figure 1).

Table 1. Patient visits in agreement or disagreement with the overall disease activity assessment.

	Remission	Low Activity	Moderate Activity	High Activity	n
Visits in agreement, no. (%)	91 (87.5)	266 (92)	80 (63.5)	39 (79.6)	476
Visits in disagreement, no. (%)	13 (12.5)	23 (8)	46 (36.5)	10 (20.4)	92
Total	104	289	126	49	568

Table 2. Variables associated with disagreement in the multivariate analysis.

Variables	Model 1		Model 2	
	OR (95% CI)	p	OR (95% CI)	p
Female gender	0.72 (0.41–1.27)	0.27	0.64 (0.37–1.10)	0.10
Age	—	NS	—	NS
Patient global disease activity	0.99 (0.98–1.01)	0.28	0.99 (0.98–1)	0.17
Swollen joint count	—	NS	—	NS
Tender joint count	1.04 (0.99–1.09)	0.06	1.05 (1–1.10)	0.02
Health Assessment Questionnaire	—	NS	1.80 (1.11–2.94)	0.02
Erythrocyte sedimentation rate	—	NS	—	NS
C-reactive protein	—	NS	—	NS
Disease activity state by consensus				
Remission	Reference	—		
Low	0.60	0.19	NI	NI
Moderate	3.77	< 0.01		
High	1.53	0.46		

NS: not significant; NI: not included.

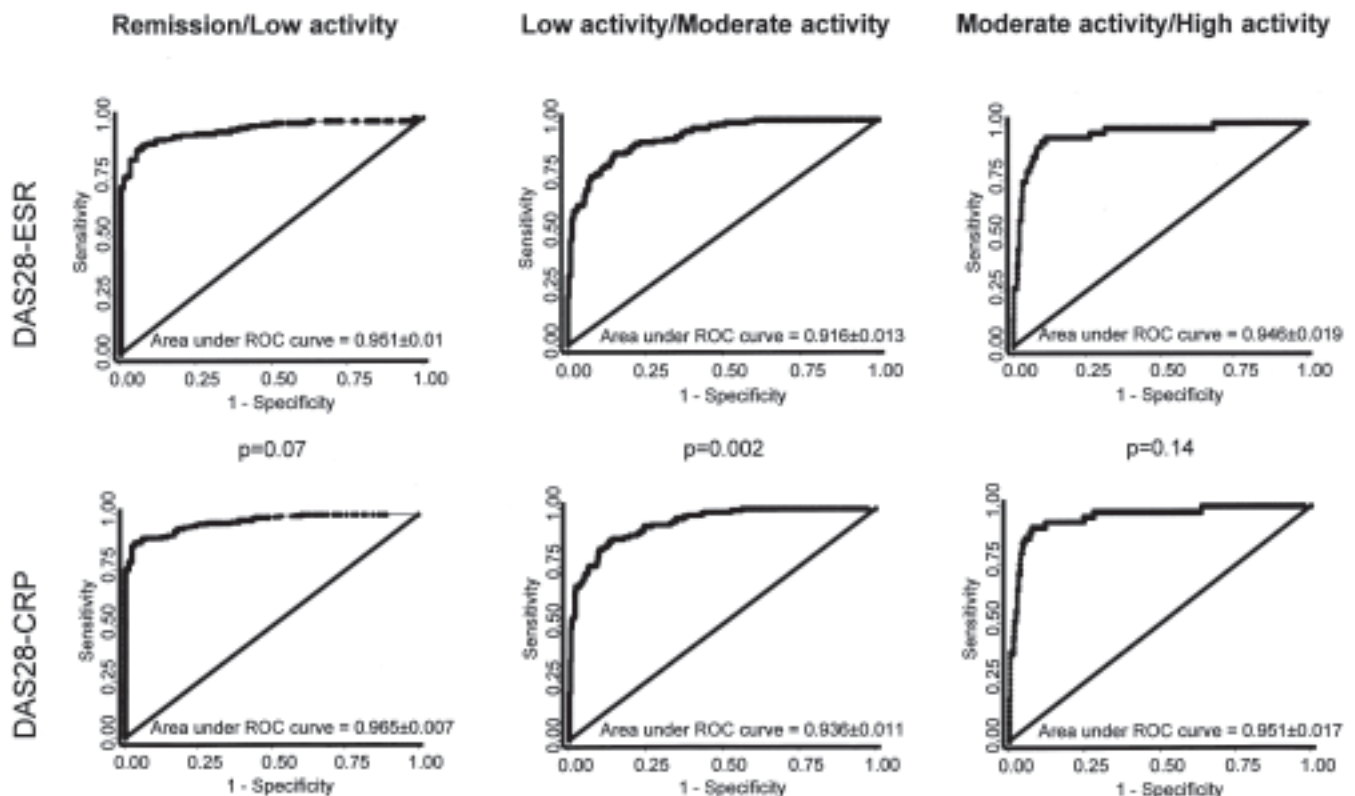


Figure 1. DAS28-CRP discriminated the disease activity state better than the DAS28-ESR. The “roccomp” command of Stata was used to establish whether there were significant differences between the areas under the receiver operating characteristic (ROC) curves of DAS28-CRP (lower panels) or DAS28-ESR (upper panels).

In addition, we estimated a set of cutoff points for DAS28-CRP according to the classification of disease activity that was based on the conventional DAS28-ESR cutoff points, as described<sup>7</sup>. Although this strategy does not allow us to estimate which index is more accurate, we obtained the following cutoff values: 2.35 for remission/low activity (Se 90%; Sp 92%), 2.95 for low/moderate activity (Se 91.5%; Sp 90.5%), and 4.35 for moderate/high activity (Se 93%; Sp 96%).

*Estimation of the MDA cutoff point for DAS28-CRP and DAS28-ESR.* A more practical approach for daily clinical practice could be to use the MDA rather than the 4 levels of disease activity. The best MDA threshold estimated for DAS28-ESR was 2.8 (Se 86%; Sp 83%), and for DAS28-CRP 2.6 (Se 85%; Sp 89%). The area under the ROC curve was again significantly higher for DAS28-CRP versus DAS28-ESR (Figure 2), suggesting that, to assess MDA, the DAS28-CRP cutoff point is more accurate.

## DISCUSSION

The DAS28-CRP was developed on the basis that CRP is a more reliable acute-phase reactant than ESR<sup>6</sup>; therefore CRP should be more useful to evaluate disease activity in patients with RA. When the formula to calculate DAS28-CRP was first described, it appeared to correlate well with DAS28-ESR (<http://www.das-score.nl>), although it has only recently been validated with respect to functional disability and radiographic progression in patients with RA<sup>13</sup>. Despite this validation, the authors suggested that it might be necessary to derive a new set of cutoff points for DAS28-CRP to increase the agreement between DAS28-CRP and DAS28-ESR<sup>13</sup>.

We show that cutoff points estimated specifically for DAS28-CRP in an early arthritis population are lower than those used classically for DAS28-ESR. This is the case when we calculated the DAS28-CRP cutoff points using the

classic gold standard proposed for DAS28-ESR<sup>4</sup>, as well as when we estimated new sets of cutoff points for both these indices using our evaluation of disease activity as the gold standard. This latter strategy was adopted mainly for 2 reasons: (1) to be able to compare “the accuracy of DAS28-ESR versus DAS28-CRP”; and (2) because the classic cutoff points calculated for DAS28-ESR are derived, through a mathematical transformation, from those estimated for the original DAS.

The original DAS was developed on the basis of prospective data in patients with recent-onset RA untreated at baseline<sup>1</sup>. Active disease was defined as a need to start or to modify DMARD therapy. Not initiating or modifying DMARD therapy over a 1-year period, or discontinuing DMARD due to disease remission, defined minimally active disease. By contrast, our gold standard to define levels of disease activity was the physicians’ assessment of the clinical data recorded in the database. Despite these differences, our estimated cutoff points for DAS28-ESR are quite similar to those calculated for DAS28-ESR from the original DAS. Further, they are even closer to those described by Aletaha, *et al*<sup>14</sup>, who also used expert assessment of patient files.

The second finding from our study is that DAS28-CRP seems more accurate than DAS28-ESR to determine RA activity in our population. This finding might be biased by our confidence in CRP versus ESR. However, we also replicated this observation through independent and validated criteria for MDA, as reported<sup>11</sup>. This outcome measurement was proposed because assessing patients with low and moderate activity is very difficult, and there is currently no precise definition of RA remission<sup>15</sup>. Therefore, Quinn, *et al* proposed that current RA treatment should aim to achieve MDA<sup>16</sup>. Interestingly, our best MDA thresholds for both DAS28-ESR and DAS28-CRP were very close to their respective cutoff values for remission, suggesting that MDA and clinical remission are very similar concepts.

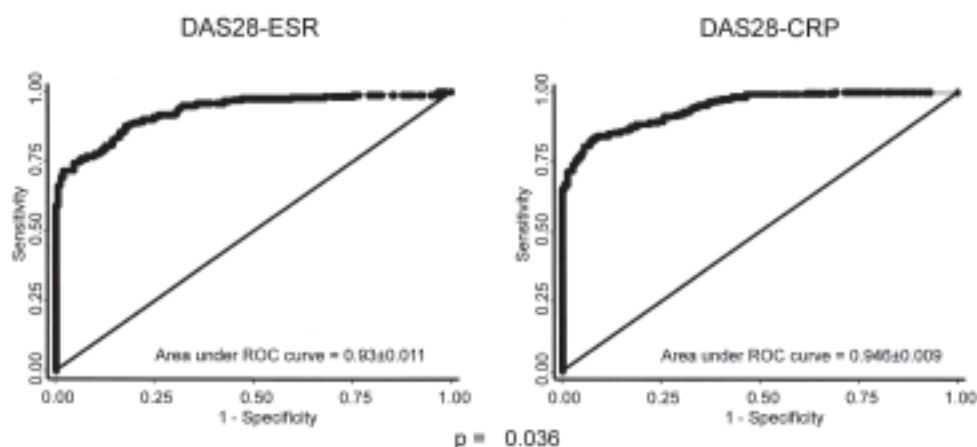


Figure 2. DAS28-CRP better discriminates the minimal disease activity versus DAS28-ESR. We used the “roccomp” command of Stata to determine whether the areas under the receiver operating characteristic (ROC) curves of DAS28-CRP (right panel) or DAS28-ESR (left panel) were significantly different.



A possible limitation of our study is that we considered patients fulfilling ACR criteria for RA as well as those with undifferentiated arthritis. Although this might introduce bias, we demonstrated previously that both DAS28-ESR and DAS28-CRP behave similarly in both these subgroups of patient<sup>8</sup>. Alternatively, our study might be biased by the fact that all the evaluators belonged to the same department. However, there was some disagreement between physicians, especially in the moderate level of disease activity. Hence, individual variation in the perception of intermediate levels of disease activity is likely to exist.

Our work supports previous studies suggesting the cutoff points for DAS28-CRP and DAS28-ESR are distinct. In addition, our data suggest that DAS28-CRP is more accurate. On the other hand, we consider that further studies are necessary to confirm whether these cutoff points for DAS28-CRP are heterogeneous in different populations. Lastly, if DAS28-CRP is confirmed to be more accurate than DAS28-ESR in other populations, specific DAS28-CRP EULAR response criteria should be described with ad-hoc cutoff values and magnitudes of improvement.

## ACKNOWLEDGMENT

We thank Loreto Carmona, MD, PhD, for the critical reading of the manuscript; and Jose M. Álvaro Gracia, MD; Inmaculada Carvajal, MD; Rosario García-Vicuña, MD, PhD; Alicia Humbría, MD, PhD; Juan P. López Bote, MD, PhD; Pedro Sabando, MD, PhD; Eva Tomero, MD; and Esther Vicente, MD, PhD, for their continuous support with the early arthritis register.

## REFERENCES

1. Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49:916-20.
2. Gonzalez-Alvaro I, Descalzo MA, Carmona L. Trends towards an improved disease state in rheumatoid arthritis over time: influence of new therapies and changes in management approach: analysis of the EMECAR cohort. *Arthritis Res Ther* 2008;10:R138.
3. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364:263-9.
4. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44-8.
5. Fransen J, Welsing PMJ, de Keijzer RMH, van Riel PLCM. Disease activity score using C-reactive protein: CRP may replace ESR in the assessment of RA disease activity [abstract]. *Ann Rheum Dis* 2003;62:s151.
6. Kushner I. C-reactive protein in rheumatology. *Arthritis Rheum* 1991;34:1065-8.
7. Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. *Ann Rheum Dis* 2007;66:407-9.
8. Castrejon I, Ortiz AM, García-Vicuña R, López-Bote JP, Humbría A, Carmona L, et al. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? *Clin Exp Rheumatol* 2008;26:769-75.
9. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.
10. Esteve-Vives J, Batlle-Gualda E, Reig A. Spanish version of the Health Assessment Questionnaire: reliability, validity and transcultural equivalency. Grupo para la Adaptacion del HAQ a la Poblacion Espanola. *J Rheumatol* 1993;20:2116-22.
11. Wells GA, Boers M, Shea B, Brooks PM, Simon LS, Strand CV, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;32:2016-24.
12. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.
13. Wells GA, Becker JC, Teng J, Dougados M, Schiff MH, Smolen JS, et al. Validation of the 28-Disease Activity Score (DAS28) and EULAR response criteria based on CRP against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on ESR. *Ann Rheum Dis* 2009;68:954-60.
14. Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states. *Arthritis Rheum* 2005;52:2625-36.
15. Shaver TS, Anderson JD, Weidensaul DN, Shahouri SS, Busch RE, Mikuls TR, et al. The problem of rheumatoid arthritis disease activity and remission in clinical practice. *J Rheumatol* 2008;35:1015-22.
16. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003;21 Suppl 31:S154-7.

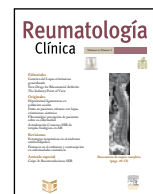
**1.3. Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid. (Reumatol Clin 2010;6:134-40)**





# Reumatología Clínica

www.reumatologiaclinica.org



## Original article

## Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid

Isabel Castrejón Fernández,<sup>a,\*</sup> Juan A. Martínez-López,<sup>b</sup> Ana M. Ortiz García,<sup>a</sup> Loreto Carmona Ortells,<sup>b</sup> Rosario García-Vicuña,<sup>b</sup> and Isidoro González-Álvarez<sup>a</sup>

<sup>a</sup>Servicio de Reumatología, Hospital Universitario "La Princesa", Madrid, Spain

<sup>b</sup>Unidad de Investigación, Fundación Española de Reumatología, Madrid, Spain

### ARTICLE INFO

#### Article history:

Received April 19, 2009

Accepted September 3, 2009

#### Keywords:

Rheumatoid arthritis

Gender

Disease Activity Indices

DAS28

ESR

CRP

### ABSTRACT

**Objective:** To evaluate the differences between the responses to treatment using DAS28 based on erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) in male and female patients. We then analyzed the individual behaviour of each component in a cohort of early arthritis patients in zone 2 of Madrid.

**Patients and methods:** We studied a total of 134 patients (77.6% women) who met the American College of Rheumatology (ACR) criteria for the diagnosis of rheumatoid arthritis (RA) belonging to an early arthritis register of the Hospital de La Princesa. We performed 4 visits following a standardized protocol which included necessary variables to calculate the DAS28 with ESR and CRP as well as determining the treatment received by the patients. We analyzed the differences in responses to treatment in males and females using both indexes, as well as their component and the assessment of the disease by the physician.

**Results:** Women had higher disease activity and disability at baseline. Although they received more intensive treatment, their average value of DAS28 remained significantly higher compared to men during the follow-up. By contrast, the global disease assessment evaluated by the patient and by the physician remained similar in both gender. When we analyze the DAS28 components separately, it was observed that this discrepancy was due mainly to the tender joints count and the ESR.

**Conclusions:** Women with early RA have higher DAS28ESR scores as a result of higher tender joint counts and ESR. This may represent bias when assessing the response to treatment using the DAS28ESR.

© 2009 Elsevier España, S.L. All rights reserved.

## Influencia del género en la respuesta al tratamiento en una cohorte de pacientes con artritis reumatoide precoz del área 2 de la Comunidad de Madrid

### RESUMEN

#### Palabras clave:

Artritis reumatoide

Género

Índices de actividad

DAS28

Velocidad de sedimentación globular

Proteína C reactiva

**Objetivo:** Valorar las diferencias de respuesta al tratamiento mediante DAS28 calculado mediante velocidad de sedimentación globular (VSG) y proteína C reactiva teniendo en cuenta el género del paciente y analizar el comportamiento individual de cada uno de sus componentes en una cohorte de pacientes de artritis precoz en el área 2 de la Comunidad de Madrid.

**Pacientes y métodos:** Se estudiaron un total de 134 pacientes (77,6% mujeres) que cumplían criterios del Colegio Americano de Reumatología para el diagnóstico de artritis reumatoide del registro de artritis precoz del Hospital de La Princesa. En dicho registro se realizaron 4 visitas protocolizadas en las que se recogen de forma sistemática los datos necesarios para calcular el DAS28 con VSG y proteína C reactiva, así como el tratamiento prescrito a los pacientes. Se analizaron las diferencias por género en la respuesta al tratamiento mediante ambos índices compuestos, así como de las variables que los componen y la valoración de la enfermedad por el médico.

\* Corresponding author.

E-mail address: isabelcastrejon@ser.es, izonci@yahoo.com (I. Castrejón Fernández).

**Resultados:** Las mujeres presentaron mayor actividad de la enfermedad y discapacidad al inicio del seguimiento. A pesar de que éstas recibieron un tratamiento más intenso, su valor promedio de DAS28 no llegó a igualarse con el de los hombres a lo largo del seguimiento. Por el contrario, la valoración de la enfermedad por parte del paciente y del médico sí llegó a igualarse. Al analizar los componentes del DAS28 por separado, se observó que esta discordancia era debida principalmente a las variables VSG y recuento de articulaciones dolorosas.

**Conclusiones:** La VSG y el recuento de articulaciones dolorosas causan un sesgo en la evaluación de la actividad de la artritis reumatoide con el DAS28 que puede afectar a la evaluación de la respuesta al tratamiento.

© 2009 Elsevier España, S.L. Todos los derechos reservados.

## Introduction

The management of rheumatoid arthritis (RA) has improved considerably in recent years. In addition to the development of new drugs, an important part of this improvement could be due to better management of classic disease-modifying drugs (FAME), including methotrexate, as well as a result of the use of composite indexes which assess the activity of disease and help us optimize treatment decisions.<sup>1,2</sup>

Currently, the DAS28 is probably the most widely used composite index in daily clinical practice. This index includes a weighted number of tender and swollen joints on a 28 joint count, evaluation of disease activity by the patient and the erythrocyte sedimentation rate (ESR) and acute-phase reactants.<sup>3</sup> However, during the past decade some limitations of this index have been demonstrated. Overall, women score higher and are therefore classified as in remission less frequently than men.<sup>4</sup> This is due, in part, to the fact that women have a higher ESR.<sup>5</sup> On the other hand, using the remission definition of the American College of Rheumatology criteria,<sup>6</sup> no gender differences are seen because it gives a uses a different cutpoint for the ESR (women <30 and men <20).

For this reason we have developed a formula for calculating the DAS28 using C-reactive protein (CRP) instead of ESR (<http://www.das-score.nl>). Although there is a good correlation between both indices, different studies have shown that they are not fully equivalent.<sup>7,8</sup> In this sense, our group has reported that the DAS28 with ESR tends to give higher values, mainly in women and in patients with longer disease progression.<sup>9</sup>

Therefore, the objective of this study is to assess the differences in response to treatment with DAS28 calculated by using ESR and CRP and taking into account the gender of the patient, as well as analyzing the individual behavior of each of its components.

## Patients and methods

We used data from patient records belonging to the recent onset arthritis clinic of the Hospital Universitario de La Princesa in area 2 of the Community of Madrid. In this clinic we received patients derived from primary care with two or more swollen joints for at least four weeks and with no more than a year of progression. We excluded patients with microcrystalline arthritis, septic arthritis, spondyloarthropathies or connective tissue diseases. The study protocol was reviewed and approved by the ethics committee and all participating patients signed an informed consent.

Registration began in September 2001 and the cutoff date for data analysis was July 2008. 484 visits of 134 patients were studied (mean=3.6 visits per patient with a range of 2 to 4 visits per patient). Patients included fulfilled the American College of Rheumatology criteria for the classification of RA at the end of follow-up.<sup>10</sup> Seventy-seven point six percent were women with an age at onset of 66 years for men and 51 years for women, and this difference was statistically significant (Table 1). Four visits were conducted according to protocol in a follow-up period of two years, consisting of a baseline visit, a visit at 6 months, one year and two years. In each visit, clinical and demographic data were collected and included in a database, such as the 28 tender and swollen joint counts, global assessment of disease by the physician (VGEM) and by the patient (VGEP) on a visual analog scale and blood tests were conducted, including ESR, CRP, rheumatoid factor, anti-CCP antibodies (ACCP) and others. Acute Phase reactants were measured by routine laboratory techniques (ESR by Westergren and CRP by nephelometry), RF by nephelometry (positive>20 IU/ml) and ACCP was determined by ELISA (Immunoscan CCPlus,® Euro-Diagnostica. Arnhem, Netherlands).

DAS28 were calculated both with CRP and ESR as previously described:

**Table 1**

Patient demographic and baseline characteristics (n=134)

	Male	Female	P	Total
Gender n (%)	30 (22%)	104 (78%)		134
Age at onset of disease, median (RI), years	66.4 (50.8–71)	51.2 (42.7–63.2)	<.01 <sup>a</sup>	54.08 (42.9–67.3)
Time since onset of disease in months (1st visit), median (RI)	5.35 (3.8–7)	6.4 (4.2–8.8)	.14	5.7 (4.2–8.5)
Positive RF, No. (%)	16 (53%)	56 (54%)	.96	72 (54%)
Positive ACCP, No. (%)	11 (39%)	54 (52%)	.22	65 (49%)
Median DAS28 (RI)	3.8 (3.3–4.9)	5 (3.9–6)	<.01 <sup>a</sup>	4.9 (3.8–5.9)
Median HAQ (RI)	0.675 (0–1.375)	1.375 (0.75–1.085)	<.01 <sup>a</sup>	1.125 (0.05–1.75)
Median ESR (RI), mm first hour	22 (11–38)	28 (18–46)	.11	24.5 (17–45)
Median CRP (RI), mg/dl	1 (0.6–1.67)	0.8 (0.3–1.8)	.30	0.8 (0.3–1.8)
Median PJC (RI)	2 (0–6)	6 (2–12)	<.01 <sup>a</sup>	5.5 (2–11)
Median SJC (RI)	4 (2–6)	5 (2–10)	.11	5 (2–9)
Median VGEM (RI)	38.5 (24–49.5)	47 (30–60)	.04 <sup>a</sup>	45 (26.5–57)
Median VGEM (RI)	31 (25–50)	47.5 (25–66)	.03 <sup>a</sup>	40.5 (25–65)

ACCP indicates anti-citrullinated peptide antibody; CRP, C reactive protein; DAS28, Disease Activity Score 28 (DAS28); ESR, erythrocyte sedimentation rate; PJC, painful joint count; RF, rheumatoid factor; RI, interquartile range; SJC, swollen joint counts; VGEM, global assesment on the part of the physician; VGEP, glibal evaluation on the part of the patient. aStatistically significant differences.

$$\begin{aligned} \text{DAS28ESR} &= 0.56 \cdot \bar{O}(\text{TJC28}) + 0.28 \cdot \bar{O}(\text{SJC28}) + 0.70 \cdot \ln(\text{ESR}) + 0.014 \cdot (\text{GDAP})^3 \\ \text{DAS28CRP} &= 0.56 \cdot \bar{O}(\text{TJC28}) + 0.28 \cdot \bar{O}(\text{SJC28}) + 0.36 \cdot \ln(\text{C-RP} + 1) + 0.014 \cdot (\text{GDAP}) + 0.96 \\ &(\text{http://www.das-score.nl}) \end{aligned}$$

In the arthritis of recent onset clinic (ARC) the following data is also systematically collected: the date of start and end of each DMARD and the maximum and minimum dose reached throughout the follow up. This allowed us to determine which DMARD was prescribed at each visit to the patient but not the dose prescribed at each visit. The database does not contemplate a treatment protocol, with decisions being made by the rheumatologist following each patient.

Statistical analysis was performed with Stata 9.2® for Windows (StataCorp LP College Station, TX, USA). We calculated the median and interquartile range of each variable, in some cases the data is shown as mean and deviation. To evaluate differences between groups, we employed the Mann Whitney U test for continuous independent variables. For qualitative variables we used the chi-square test. A  $P < .05$  was considered statistically significant.

## Results

### Patient characteristics at first visit

Baseline data of patients is reflected in Table 1. Before initiating treatment, women showed greater activity of the disease and greater disability, as reflected by higher values of DAS28 and HAQ. The ESR showed a tendency to be higher in women. However, there was no gender difference in the percentage of patients with severity markers such as rheumatoid factor and ACCP, or duration of disease at first visit (Table 1).

### Treatments

As a result of their greater degree of disease activity, women received more aggressive treatment. So the percentage of women who did not receive DMARDs during follow up was reduced with respect to that of men, while the percentage of women treated with combination therapy was higher, and these differences were statistically significant (Table 2). With regard to steroids, there were no significant differences by gender in the subsequent visits (data not shown).

Regarding the type of DMARD used in men and women, the most frequently used were methotrexate and antimalarials with similar usage rates in both genders, but women received significantly higher doses of methotrexate than men. Furthermore, the percentage of women to which leflunomide was prescribed was twice as many as men in the number of total visits and this difference was statistically significant ( $P = .026$ , Table 3). Also, TNF-blocking agents were used more often in women, although this difference was not statistically significant (Table 3).

### Disease activity

Although women received more aggressive treatment since the onset of follow up, the activity level measured by DAS28 with ESR never matched that of men (Figure 1A). When we analyzed the behavior of DAS28 calculated with CRP, differences with men, once treatment was established, were not as striking, but persisted during follow-up and were not statistically significant except at the 1-year visit when it was close to being statistically significant (Figure 1B). One possibility is that despite treatment, women had responded less and maintain a higher activity of the disease and, therefore, maintain higher values on the indices. To test this possibility we analyzed the VGEM VGEP throughout follow up in both sexes. As shown in

**Table 2**  
Number of patients without treatment, in monotherapy or as combined therapy in visits 2, 3, and 4

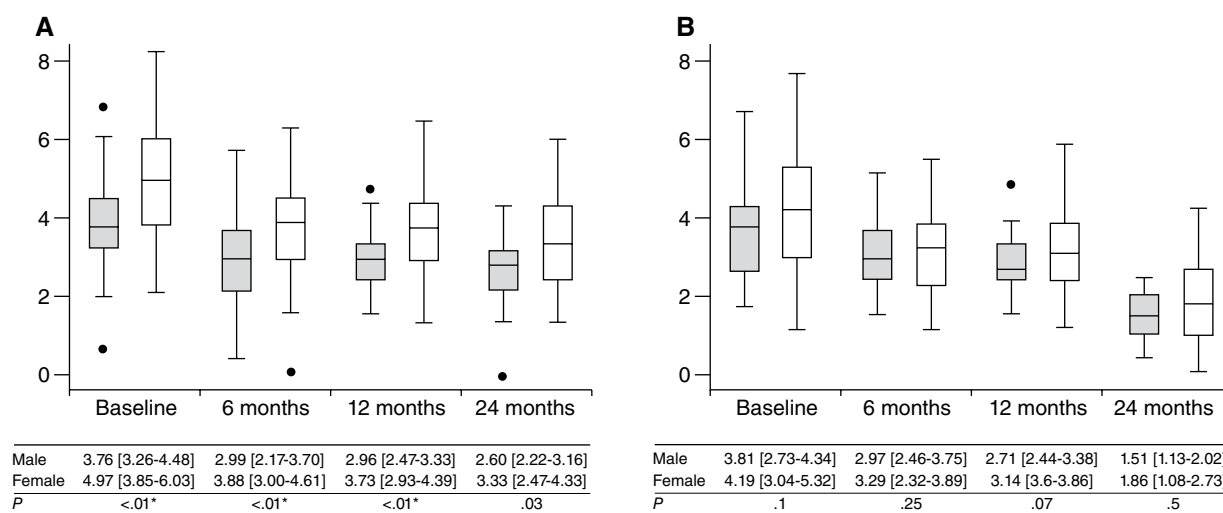
Treatment	Visit 2 n=108		Visit 3 n=118		Visit 4 n=124	
	Male	Female	Male	Female	Male	Female
No treatment	4 (17%)	8 (9%)	4 (15%)	1 (1%)	4 (15%)	6 (6%)
Monotherapy	17 (74%)	56 (66%)	20 (77%)	59 (64%)	19 (70%)	49 (51%)
Combined	2 (9%)	21 (25%)	2 (8%)	32 (35%)	4 (15%)	42 (43%)
P	.18		<.001 <sup>a</sup>		.018 <sup>a</sup>	

<sup>a</sup>Statistically significant.

**Table 3**  
Use of DMARD in men and women in the total number of visits. The second column reflects the duration of treatment in days and the third column reflects the maximum dose reached for each one of the treatments

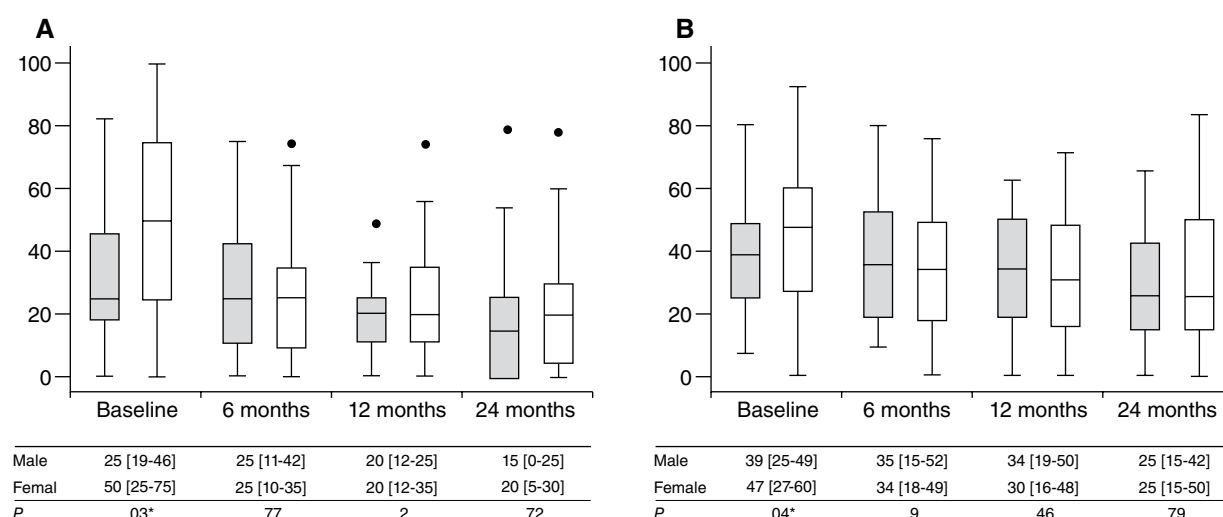
Treatment	Number of visits with DMARD during follow up n (%)			Time of treatment, days Mean±SD			Maximum dose, mg Median (RI)		
	Female	Male	P	Female	Male	P	Female	Male	P
Methotrexate	218 (76)	58 (71)	n.s.	569±29	487±65	n.s.	17.5 (15-20)	15 (10-17.5)	<.001
Antimalarials	69 (25)	15 (19)	n.s.	211±35	87±42	.03	400 (200-400)	300 (200-400)	n.s.
Leflunomide	62 (22)	8 (11)	.026	150±24	47±30	.02	14.5 (10-20)	20 (10-20)	n.s.
Sulphasalazine	13 (5)	3 (4)	n.s.	29±11	30±30	n.s.	1,500 (1,500-1,500)	1,500 (1,500-1,500)	n.s.
Gold salts	4 (1)	4 (4)	n.s.	10±7	37±27	n.s.	200 (200-200)	125 (50-200)	n.s.
Anti-TNF	11 (4)	0 (0)	.079	27±9	0	.07	n.d.	n.d.	–

SD indicates standard deviation.



**Figure 1.** DAS28 values in a population of recent onset Rheumatoid Arthritis calculated with: 1A) ESR and 1B) CRP. Data is presented as median (line within the box, 75 and 25 percentiles [superior and inferior limits of the box, respectively], as well as percentiles 95 and 5 [superior and inferior lines outside the box]). Data from males is presented in gray and women in white. Data from tables are presented as median and the interquartile range.

\*Statistically significant differences with the Mann Whitney U test.



**Figure 2.** Global evaluation of the disease: A) performed by the physician and B) performed by the patient in a population of recent onset arthritis during the first four follow up visits. At the beginning of follow up, both the evaluation of the disease by the physician as well as by the patient is greater in women but in both cases tend to equalize during follow up. Data is presented as a median (line within the box), percentiles 75 and 25 (superior and inferior limits of the box, respectively), as well as percentiles 95 and 5 (superior and inferior lines outside the box). Data belonging to males is shown in gray and females in white. Data of the tables are presented as median and interquartile range.

\*Statistically significant differences with the Mann Whitney U test.

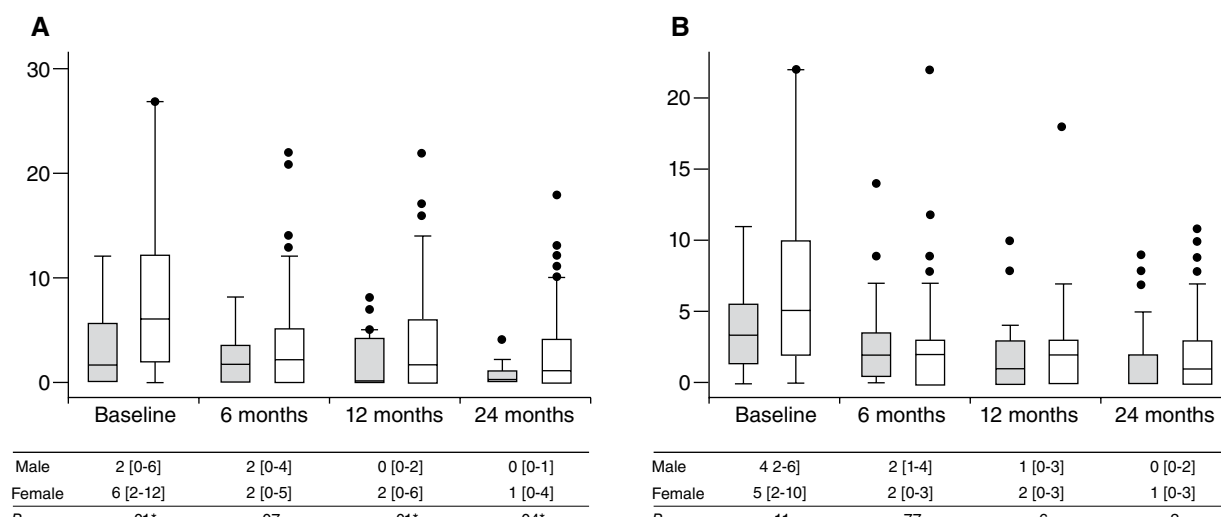
Figure 2, although both VGEM and VGEP were higher in women before treatment, these evaluations were similar in both genders in subsequent visits.

Given this discrepancy, we analyzed the behavior of the various components of DAS28, noting that women have significantly higher tender joint counts over follow up (Figure 3A). By contrast, even at the start of follow up, the swollen joint count was higher in women, this variable matched that of men after the start of treatment (Figure 3B). Moreover, as is well known, women had higher ESR values throughout the follow-up, while CRP showed no statistically significant differences (Figure 4A and B, respectively).

## Discussion

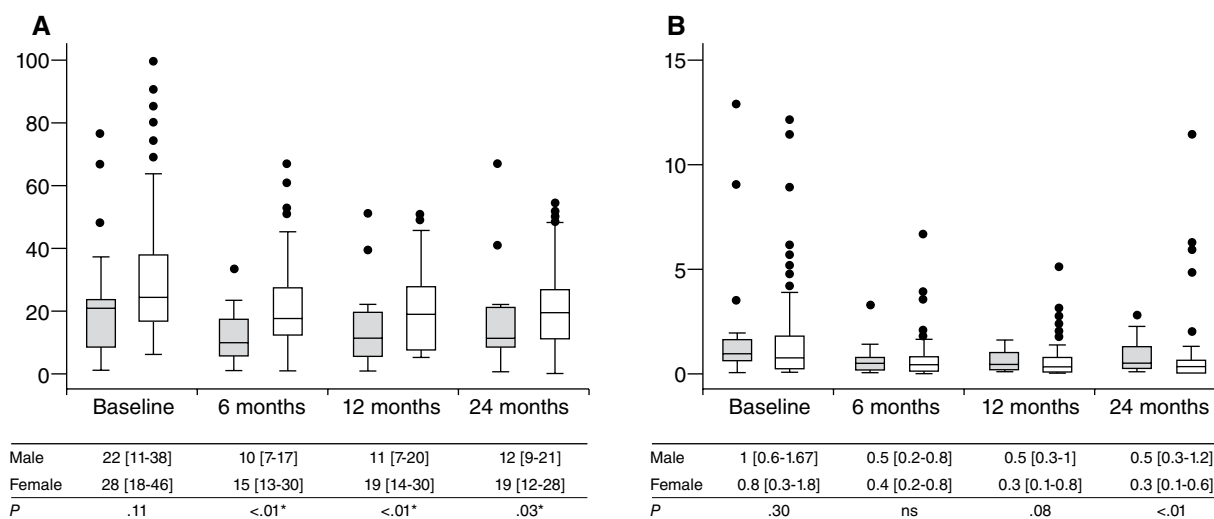
The main finding of our study is that there is a gender difference in treatment response as assessed by DAS28, as well as in the assessment of the disease by the doctor and the patient. The main differences responsible for this are the ESR in the classical DAS28, and the tender joint count in the classic-DAS28 and the one calculated using CRP.

These differences in the DAS28 due to gender have been described by other authors. Several studies have been seen in which women tended to score higher than men in this index,<sup>11</sup> mainly because



**Figure 3.** A) tender joint count B) swollen joint counts in subsequent visits in a recent onset arthritis cohort. The painful joint count is higher during practically all of the follow up, while the swollen joint count is higher in women during their first visit (not statistically significant) but tends to equalize once treatment was started. Data are presented as median (line within the box, percentiles 75 and 25 [superior and inferior limits of the box, respectively], as well as percentiles 95 and 5 [superior e inferior lines outside the box]). Data from males is shown in gray and females in white. Data from the tables is presented as median and interquartile range.

\*Statistically significant differences with the Mann Whitney U Test.



**Figure 4.** A) ESR and B) CRp values during follow up. ESR has higher values in women in all of the visits while CRP showed this only at the beginning of follow up. Data is presented as median (line within the box, percentile 75 and 25 [superior and inferior limits if the box, respectively], as well as percentiles 95 and 5 [superior and inferior line outside the box]). Data from males is shown in gray and women in white. Data from tables is presented as males and interquartile range.

\*Statistically significant differences with the Mann Whitney U test.

women have higher levels of ESR. But even if the ESR is a clear cause for this difference, the use of CRP does not fully solve the differences by gender, and therefore other factors of the DAS28 should contribute to these differences.

Our data provides evidence that another factor contributes to these differences, such as the painful joint count and that, despite more intensive therapy, women on average had higher tender joint counts than men throughout follow-up. Some studies have described how the perception of pain is more pronounced in women than in

men with RA,<sup>12</sup> and this fact has also been confirmed in the general population.<sup>13</sup> The physiological explanation for this gender difference could be that women have greater activation of nociceptive unmyelinated C type fibers,<sup>14,15</sup> and have a decreased response to analgesics acting on opioid receptors.<sup>16</sup> Therefore, although some authors consider that tender joint counts should have an important weight in the assessment of disease activity,<sup>17</sup> it is possible that this variable reflects a situation external to RA and, therefore, represent bias in the assessment of disease activity.

One could argue against our results that both VGEF and VGEM are an unsound 'gold standard' and that the highest tender joint counts in women themselves are a manifestation of the increased activity of RA. However, other authors have also found gender-related differences in the DAS28 without it affecting a variable with greater weight such as radiological progression after 5 years of follow-up.<sup>18</sup> This fact, together with our data, suggests that a higher DAS28 does not always necessarily mean a greater aggressiveness and a worse outcome of disease in women.

The impact of this gender difference in the assessment of the disease by DAS28 is very important because of its use in clinical trials to assess response to treatment.<sup>19</sup> On the one hand, by using the DAS28 as an assessment tool, women would reach remission less frequently than men.<sup>4</sup> On the other hand, since women start with higher DAS28 levels at the onset of follow up, they would have worst rates of treatment response according to the EULAR criteria.<sup>20</sup> From a clinical point of view, another problem that arises is that many clinical practice guidelines suggest a DAS28 cutoff point after which biological therapy should be considered.<sup>21</sup> In this way, it would be easier to start this type of treatment in female patients.

This leads us to consider whether to continue using the classic DAS28 in the evaluation of patients with RA. As demonstrated in this study, DAS28 calculated with CRP has little bias in the assessment of disease activity, although certain gender differences persist due to the high weight of the tender joint count.

With regard to other indices available, SDAI and CDAI also prevent part of the bias that occurs as they include the CRP as an acute phase reactant.<sup>22,23</sup> However, these indices have some drawbacks; first, they do not ponder the different variables that constitute them and, in addition, CRP is also included as an absolute value despite not having a normal distribution, which can represent a problem. In our population, SDAI also shows differences by gender (Figure 5), with a tendency to be higher in women, although the only statistically significant difference was seen in the visit at one year.

In summary, our study shows that the evaluation of treatment response rates based on currently available indices has a bias related to gender differences. It would be important to develop new

indices to avoid this bias and therefore be more objective when making treatment decisions and evaluating the results in clinical trials.

### Conflict of interest

Dr. Isidoro González-Alvaro has received research funding from Abbott Laboratories, Sanofi-Aventis and Bristol-Myers Squibb during the last five years. All these research projects are unrelated to this work.

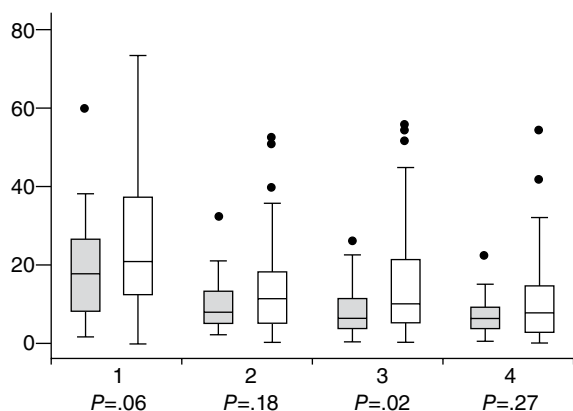
### Financing

This work was supported in part by the RETICS Program, RD08/0075 (Riera) Institute of Health Carlos III (ISCIII) and by projects FIS 05/2044 and No. 08/0754, and a grant for stimulatory research work awarded to Dr. Isidoro González Álvaro, promoted by the Carlos III Health Institute (ISCIII).

The work of Dr. Elizabeth Castrejón has been funded in part by a grant from the Serap project of the Spanish Foundation of Rheumatology. Dr. JA Martínez has a research training contract ISCIII Rio Hortega.

### References

- González-Álvarez I, Descalzo MA, Carmona L. Trends towards an improved disease state in rheumatoid arthritis over time: influence of new therapies and changes in management approach: analysis of the EMECAR cohort. *Arthritis Res Ther*. 2008;10:R138.
- Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet*. 2004;364:263-9.
- Prevoo ML, Van 't Hof MA, Kuper HH, Van Leeuwen MA, Van de Putte LB, Van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum*. 1995;38:44-8.
- Forslind K, Hafstrom I, Ahlmen M, Svensson B. Sex: a major predictor of remission in early rheumatoid arthritis? *Ann Rheum Dis*. 2007;66:46-52.
- Miller A, Green M, Robinson D. Simple rule for calculating normal erythrocyte sedimentation rate. *Br Med J (Clin Res Ed)*. 1983;286:266.
- Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*. 1995;38:727-35.
- Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. *Ann Rheum Dis*. 2007;66:407-9.
- Matsui T, Kuga Y, Kaneko A, Nishino J, Eto Y, Chiba N, et al. Disease Activity Score 28 (DAS28) using C-reactive protein underestimates disease activity and overestimates EULAR response criteria compared with DAS28 using erythrocyte sedimentation rate in a large observational cohort of rheumatoid arthritis patients in Japan. *Ann Rheum Dis*. 2007;66:1221-6.
- Castrejón I, Ortiz AM, García-Vicuna R, Lopez-Bote JP, Humbria A, Carmona L, et al. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? *Clin Exp Rheumatol*. 2008;26:769-75.
- Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum*. 1988;31:315-24.
- Hallert E, Thyberg I, Hass U, Skargren E, Skogh T. Comparison between women and men with recent onset rheumatoid arthritis of disease activity and functional ability over two years (the TIRA project). *Ann Rheum Dis*. 2003;62:667-70.
- Leeb BF, Haindl PM, Maktari A, Nothnagl T, Rintelen B. Disease activity score-28 values differ considerably depending on patient's pain perception and sex. *J Rheumatol*. 2007;34:2382-7.
- Unruh AM. Gender variations in clinical pain experience. *Pain*. 1996;65:123-67.
- Feine JS, Bushnell MC, Miron D, Duncan GH. Sex differences in the perception of noxious heat stimuli. *Pain*. 1991;44:255-62.
- Lautenbacher S, Rollman GB. Sex differences in responsiveness to painful and non-painful stimuli are dependent upon the stimulation method. *Pain*. 1993;53:255-64.
- Gear RW, Gordon NC, Heller PH, Paul S, Miaszkowski C, Levine JD. Gender difference in analgesic response to the kappa-opioid pentazocine. *Neurosci Lett*. 1996;205:207-9.
- Wolfe F, Pincus T, Thompson AK, Doyle J. The assessment of rheumatoid arthritis and the acceptability of self-report questionnaires in clinical practice. *Arthritis Rheum*. 2003;49:59-63.
- Ahlmen M, Svensson B, Albertsson K, Forslind K, Hafström I; BARFOT Study Group. Influence of gender on assessments of disease activity and function in early



**Figure 5.** DAI values in male and female patients during follow-up. Data is presented as median (line within the box, percentile 75 and 25 [superior and inferior limits of the box, respectively], as well as percentiles 95 and 5 [superior and inferior limits outside the box]). Data from males is shown in gray and females in white.

- rheumatoid arthritis in relation to radiographic joint damage. *Ann Rheum Dis.* 2001;69(1):230–3.
19. Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Arthritis Rheum.* 2008;59:1371–7.
20. Prevoo ML, Van Gestel AM, Van THMA, Van Rijswijk MH, Van de Putte LB, Van Riel PL. Remission in a prospective study of patients with rheumatoid arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score. *Br J Rheumatol.* 1996;35:1101–5.
21. Emery P, Van Vollenhoven R, Ostergaard M, Choy E, Combe B, Graninger W, et al. Guidelines for initiation of anti-tumour necrosis factor therapy in rheumatoid arthritis: similarities and differences across Europe. *Ann Rheum Dis.* 2009;68:456–9.
22. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23:S100–8.
23. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford).* 2003;42:244–57.



## **Chapter 2: Recommendations for the investigation and follow-up of undifferentiated arthritis**

As it was previously pointed out, the earlier disease activity is evaluated for a tight control with an appropriate intensification of therapy, the better the outcomes will be compared to a traditional approach. Furthermore, the importance of early diagnosis is increasingly relevant, as recent-onset disease is more sensitive to treatment than later-stage disease. For all these reason it would be reasonable to have validated instruments to be used in early populations. To identify validated instruments in early populations we reviewed the literature. Then, specific recommendations to evaluate early arthritis patients were formulated.

The following articles are presented:

1. Clinical composite measures of disease activity for diagnosis and follow-up of undifferentiated peripheral inflammatory arthritis: a systematic review. (J Rheumatol Suppl 2011;87:48-53)
2. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. (Ann Rheum Dis 2011;70:15-24)



**2.1. Clinical composite measures of disease activity for diagnosis and follow-up of undifferentiated peripheral inflammatory arthritis: a systematic review. (J Rheumatol Suppl 2011;87:48-53)**

# Clinical Composite Measures of Disease Activity for Diagnosis and Followup of Undifferentiated Peripheral Inflammatory Arthritis: A Systematic Review

ISABEL CASTREJÓN, LUCÍA SILVA-FERNÁNDEZ, CLAIRE BOMBARDIER, and LORETO CARMONA

**ABSTRACT. Objective.** To critically appraise the validity of activity indices used in the followup of patients with undifferentiated peripheral inflammatory arthritis (UPIA).

**Methods.** A systematic review was performed in Medline, Embase, the Cochrane Library, and abstracts presented at the 2007 and 2008 meetings of the American College of Rheumatology and European League Against Rheumatism. Selection criteria were: patients with UPIA, the assessment of instruments to evaluate disease activity, and assessment of validity of the instruments. Two reviewers screened titles and abstracts independently and collected data using ad hoc standard forms.

**Results.** The search yielded 179 articles and 834 abstracts, of which 4 articles and 1 abstract were included. We found no study that validated Disease Activity Score (DAS), Clinical Disease Activity Index (CDAI), or Simplified Disease Activity Index (SDAI). Included studies addressed validation of 4 questionnaires: WHO Disability Assessment Schedule (WHODAS), London Handicap Scale (LHS), Disease Repercussion Profile (DRP), and the Health Assessment Questionnaire (HAQ); and 3 indexes: RA Disease Activity Index (RADAI), McGill Range of Motion Index (McROMI), and NOAR Damaged Joint Count (NOAR-DJC). Questionnaires were self-administered and feasible; RADAI was the most feasible index. Internal consistency was studied in the questionnaires (Cronbach's  $\alpha > 0.83$ ). Responsiveness was tested in the DRP, LHS, and HAQ, but the approach to study sensitivity to change was poorly explained, with no clear intervention. Construct validity, examined by means of convergence with other instruments, was generally moderate, and slightly higher for the RADAI.

**Conclusion.** No instrument of disease activity has been fully validated for use in UPIA. We found no direct evidence of what is the most useful index to follow up patients with UPIA. (J Rheumatol 2011;38 Suppl 87:48–53; doi:10.3899/jrheum.101075)

## Key Indexing Terms:

UNDIFFERENTIATED ARTHRITIS  
OUTCOME MEASURES

ASSESSMENT TOOLS

DISEASE ACTIVITY  
SYSTEMATIC REVIEW

*From the Rheumatology Department, Hospital Universitario la Princesa; Rheumatology Department, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain; Division of Rheumatology and Department of Health Policy, Management, and Evaluation, University of Toronto; Division of Clinical Decision Making and Health Care, Toronto General Research Institute, University Health Network, Mount Sinai Hospital, Toronto, Canada; and Research Unit, Fundación Española de Reumatología, Madrid, Spain.*

*Supported by an unrestricted educational grant from Abbott. Abbott had no role in the design, literature search, data collection, data analysis, data interpretation, or writing of this report. Dr. Bombardier holds a Pfizer Chair and Canada Research Chair in Knowledge Transfer for Musculoskeletal Care.*

*I. Castrejón, MD, Rheumatology Department, Hospital Universitario la Princesa; L. Silva-Fernández, MD, Rheumatology Department, Hospital Universitario Puerta de Hierro Majadahonda; C. Bombardier, MD, FRCP, Division of Rheumatology and Department of Health Policy, Management, and Evaluation, University of Toronto; Division of Clinical Decision Making and Health Care, Toronto General Research Institute, University Health Network, Mount Sinai Hospital; L. Carmona, MD, PhD, Research Unit, Fundación Española de Reumatología.*

*Address correspondence to Dr. I. Castrejón, S. Reumatología, Hospital La Princesa, C/ Diego de León 62, 28006 Madrid, Spain.  
E-mail: isabelcastrejon@ser.es*

Many instruments for disease activity assessment have been developed in recent years. These indices are frequently used in clinical trials as well as in daily practice as they are useful to evaluate response to treatment or to make a decision to start or change treatment. Use of such indices has become an important aspect of the care for patients with rheumatoid arthritis (RA)<sup>1,2</sup>. However, we are unaware if they are equally useful for patients with undifferentiated peripheral inflammatory arthritis (UPIA).

UPIA is a form of arthritis that does not fulfill classification criteria for a more definitive diagnosis. Patients with UPIA are hard to follow in clinical practice, as they comprise a very heterogeneous group, sharing characteristics of different diagnoses. Due to the lack of a more precise clinical picture and outcome, it is important to have comprehensive tools that help the clinician anticipate outcomes, including more precise diagnosis, and thus make therapeutic decisions. Studies focused on UPIA have used many different indices to evaluate outcome. However, the sole fact of using

an index in a study does not confer validity for evaluating outcome in that particular population, in this case, disease activity in UPIA. An instrument should demonstrate that it measures what is intended, discriminates between different disease states, and shows change in the numerical result when the patient improves.

Our objective was to analyze the validity of any available activity index, instrument, or scale, used to evaluate disease activity of patients with UPIA. The clinical correlate to our objective was to answer the question, “Which clinical assessments of disease activity [e.g., Disease Activity Score (DAS), Simplified Disease Activity Index (SDAI), and Clinical Disease Activity Index (CDAI)] should be done (at baseline and repeat at what interval) in patients with undifferentiated arthritis?”

## MATERIALS AND METHODS

This systematic review is part of the 3e (evidence, expertise, exchange) Initiative in Rheumatology. The 3e Initiative is a multinational effort, aimed at promoting evidence-based medicine, by formulating detailed recommendations addressing clinical problems<sup>3,4</sup>. In contrast to guidelines developed by a limited panel of experts, the 3e Initiative involves a broad international panel of rheumatologists. Further, the initiative promotes epidemiology, by teaching and conducting systematic literature research following a strict methodology<sup>5</sup>. The objective of the 3e Initiative of 2008-2009 was to develop practical recommendations for the investigation and followup of UPIA, by integrating systematically generated evidence and expert opinion of a broad panel of international rheumatologists.

*Rephrasing the research question.* The clinical question as formulated by experts from 18 countries was translated into an epidemiological research question according to the PICO (Patient, Intervention, Comparator, and Outcome) approach<sup>6</sup>. Definitions: Patients were defined as adults with UPIA; Intervention as any index, instrument, or scale used to evaluate disease activity; Comparators were the above indices compared to themselves or to another index; and Outcome was any aspect of validity: construct validity, feasibility, reliability, or responsiveness.

The final search question was rephrased as, “What are the most suitable clinical measures to evaluate the diagnosis of UPIA? and, What are the most useful indexes to evaluate the followup?”. From a clinical point of view, we wanted to know whether any index could help differentiate patients with UPIA from others that would develop a specific diagnosis.

*Scenarios.* Whether a disease activity measure is useful is a difficult question to address. We identified at least 3 possible approaches to the answer, each determining a different search strategy. The first approach would be to retrieve all studies that included common indices of disease activity (DAS, SDAI, CDAI, etc.) in which at least one group had UPIA. Next, we would evaluate whether these indices were discriminating groups of high and low activity in the UPIA arm. The second approach would be to search for all UPIA studies, and check which indices were used. This approach would tell us only if the indices discriminate between patient states. These 2 approaches would require in-depth knowledge of the studies through contact with the authors, as the information on discrimination of variables is not usually available in studies that do not specifically address validity.

The third and most objective approach, which was selected by the team, involved searching for studies in which the population studied was UPIA, in which indices were used, and that specifically measured any aspect of validity.

*Systematic literature search.* We performed a systematic literature search for articles published between 1950 and January 2009 in Medline, Embase, and the Cochrane Library, using a comprehensive search strategy (see online appendix, available from: [www.3eupia.com](http://www.3eupia.com)). Abstracts presented at

the 2007 and 2008 meetings of the European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) were also searched. Two reviewers (LS and IC) screened the titles and abstracts for selection criteria independently, using a third reviewer for consensus, and collected the data using ad hoc standard forms. Review articles were also retrieved for identifying additional references by hand search.

*Selection of articles.* Relevant articles were selected in a systematic procedure. First, titles and abstracts of all identified references were screened, excluding articles that clearly did not address the topic of interest. Second, selected articles were reviewed in the full report, applying the following inclusion criteria: validation studies, UPIA patients, adults (> 18 years), and disease activity measures. Articles that did not fulfill all the inclusion criteria were excluded (Table 1). We did not include other standard clinical monitoring measures such as pain, global assessment, joint count, erythrocyte sedimentation rate, etc., because they were being investigated in other reviews in the 3e Initiative.

*Data extraction and analysis.* Publication details, patient characteristics, instruments to measure disease activity, and aspects of the assessment of validity were extracted using standard forms. To evaluate validity of each instrument, feasibility, reliability, responsiveness, and construct validity were analyzed. To measure feasibility, we deduced from study information the time required to complete an instrument and its ease of use. We created an ad hoc measure that went from 0 (unfeasible) to 3 (completely feasible).

Reliability embraces the concept that repeat administration of a measurement tool in stable subjects will yield the same result, thus measuring an instrument's stability. Reliability also includes an instrument's internal consistency or “good construction,” as expressed by the Cronbach's alpha statistic (< 0.70 = individual items provide an inadequate contribution to the overall scale; > 0.90 suggests redundancy). The stability of the instrument should be tested twice: by the same operator at different times (intraobserver test-retest) and by a different operator (interobserver reliability) at the same time. Intra- and interobserver reliability are measured either with the kappa statistic or with the intraclass correlation coefficient.

Responsiveness, also called sensitivity to change, is defined as “the ability of an instrument to accurately detect change when it has occurred”<sup>7</sup>. It measures whether the instrument detects that the patient has improved or worsened. Responsiveness implies that an intervention with an effect of known direction is given to the studied patients. It is quantified by the effect size (ES) or the standardized response mean (SRM). In accord with the literature, ES were considered as follows: ~0.2 = small, ~0.5 = moderate, and > 0.8 = large.

Construct validity is a measure of how close to what the instrument says it measures it really measures. One way of conceptualizing construct validity is to consider it as testing hypotheses of what a valid instrument would and would not correlate with; thus the instrument is compared to other instruments measuring the same construct (high correlation) and different construct (low correlation). Establishing the validity of an instrument to measure disease activity is difficult, as no established “gold standard” is available. In most of the studies retrieved the construct validity was examined in terms of convergence with variables that should have a converging relationship (correlation > 0.60 is considered a good correlation).

## RESULTS

A total of 179 references and 834 meeting abstracts were identified. After title/abstract screening, 19 articles were retrieved for full article review, of which 4 fulfilled inclusion criteria. One meeting abstract was also included. Thus, 5 records were included addressing some aspects of the validation of 4 questionnaires and 3 physical measures. We found no study on the validation of the most common activity measures such as DAS<sup>8</sup> or SDAI<sup>9</sup> in patients with UPIA.

Table 1. Excluded studies and reason for exclusion.

Study	Reason for Exclusion
Tully <sup>21</sup>	Arthritis and osteoarthritis population, age > 60 yrs
Cohen <sup>22</sup>	RA > 1 yr duration and prognostic factors of quality of life after 5 yrs of followup
El Miedany <sup>23</sup> , Kievit <sup>24</sup> , Suurmeijer <sup>25</sup> , Salaffi <sup>26</sup>	Population was RA, not UPIA
Lerner <sup>27</sup>	RA patients as controls to validate disability in other diseases
Smolen <sup>28</sup>	RA population. Review
Hamilton <sup>29</sup>	RA population (per ACR criteria) > 1 yr disease duration. Not an index: Gait analysis on contact-sensitive walk mat system
Cole <sup>30</sup>	Scleroderma and RA population. Study goal was to examine structural validity of HAQ in patients with SSc
Saraux <sup>31</sup> , van der Helm <sup>32, 33</sup> , El Miedany <sup>34</sup>	Prediction rules (multivariate) for RA

A summary of the results of the validity of the different instruments can be found in Table 2.

*Description and feasibility of the questionnaires.* The 4 questionnaires for which a validation in a UPIA population was published were the World Health Organization Disability Assessment Schedule (WHODAS), London Handicap Scale (LHS), Disease Repercussion Profile (DRP), and the Health Assessment Questionnaire (HAQ), all self-administered.

The WHODAS is a short-form questionnaire comprising 36 Likert-formatted questions divided into 6 domains (understanding/communicating, getting around, self-care, getting along with people, life activities, and participation in society)<sup>10</sup>; final score ranges from 0 (best) to 100 (worst). The LHS has 6 domains covering handicap dimensions: mobility, physical independence, occupation, social interaction, orientation, and economic self-sufficiency; score ranges from 100 (no disadvantage) to 0 (extreme disadvantage)<sup>11</sup>. The DRP consists of 6 visual analog scales on the importance to the patients of 6 domains: functional and social activity, employment/money, relationships, emotions, and body image; score ranges from 0 (none) to 10 (extremely important)<sup>12</sup>. The HAQ includes 20 items on ability for daily activities: dressing and grooming, rising, eating, walk-

ing, hygiene, functional reach and grip and activities; score ranges from 0 (none) to 3 (complete disability)<sup>13</sup>.

Since all questionnaires are self-administered, they are feasible, although the WHODAS and the LHS seem to take longer.

*Description and feasibility of the indexes.* We identified 3 indices that had been validated in UPIA, or at least some aspect of the index had been validated: the Rheumatoid Arthritis Disease Activity Index (RADAI), McGill Range of Motion Index (McROMI), and NOAR Damaged Joint Count (NOAR-DJC).

The RADAI is a self-administered questionnaire that yields an index of activity. It comprises 5 individual items that have a high association with clinically assessed joint synovitis and acute-phase response, providing a global score<sup>14</sup>. A validity study in UPIA was available only in abstract form<sup>15</sup>.

The McROMI index is based on a visual estimate of range of motion (ROM) of 19 movements in 9 joint areas (neck, shoulder, elbow, wrist, forearm, hand, hip, knee, and ankle), all bilaterally, except for the neck. The authors propose that a limited ROM from inflammation and pain may occur early in the disease process. To obtain a score, each movement is graded from 0 to 3, 3 being the most abnormal,

Table 2. Summary of results of validity of instruments retrieved in the search strategy. Feasibility (0 = not feasible to 3 = completely feasible) is based on the time it takes, and how easy it is to answer the questions.

	Feasibility	Internal Consistency	Test-Retest	Responsiveness	Construct Validity
WHODAS	2	+++	+++		++
DRP	3	++		–	++
LHS	2	++		–	++
HAQ	3	++		–	
McROMI	1				++
NOAR-DJC	1		++		++
RADAI	3				+++

WHODAS: WHO Disability Assessment Schedule; DRP: Disease Repercussion Profile; LHS: London Handicap Scale, HAQ: Health Assessment Questionnaire; McROMI: McGill Range of Motion Index; NOAR-DJC: NOAR Damaged Joint Count, RADAI: Rheumatoid Arthritis Disease Activity Index.

and the maximum score 111. The McROMI requires assistance to complete, and the movements and scores assigned to the different degrees of mobility are not easy to remember<sup>16</sup>.

Last, the NOAR-DJC index assesses the presence or absence of deformity in 51 joints. Distal interphalangeal joints are included because the NOAR-DJC was intended for use in patients with early inflammatory polyarthritis. Deformity is defined as the inability to adopt an anatomical position and a reduction in range of movement and/or surgical alteration of the joint. The authors consider deformity to be a reversible feature of early arthritis. NOAR-DJC requires a guideline and a manikin to perform<sup>17</sup>.

Concerning feasibility of the above indexes, the RADAI seems to be the most feasible because the others are time-consuming and difficult to perform.

**Reliability and responsiveness.** The internal consistency of the 4 questionnaires was very good, all showing Cronbach's  $\alpha > 0.8$ . Regarding the indices, internal consistency is not absolutely necessary because indices are composed of very similar items. The RADAI is a mixture of a questionnaire and an index, as the way questions are asked (internal consistency) is not all that relevant. That is why we cannot say it is totally unnecessary.

On the other hand, test-retest reliability is very important for measures that imply an operator, such as mobility and deformity indices. In testing the NOAR-DJC, the interclass correlation coefficient (ICC) was slightly higher for the intraobserver study (0.88, 95% CI 0.79–0.94,  $p < 0.001$ ) than for the interobserver study (0.74, 95% CI 0.53–0.86,  $p < 0.001$ ). Both results showed fairly good reliability. Surprisingly, this was not tested in the McROMI as might have been expected in this type of index. In questionnaires, test-retest reliability was evaluated only in the WHODAS, where intraobserver reliability was assessed in 20 subjects of the population: the ICC between time 1 and time 2 was 0.94 (95% CI 0.86–0.98).

Regarding responsiveness, it was tested only in the DRP, LHS, and HAQ, all in the same study<sup>18</sup>. The approach to measure responsiveness in this study was rather poor, with no clear intervention or anticipated change.

**Construct validity.** To evaluate construct validity of the questionnaires we had to consider that the authors wanted to measure disability. The WHODAS has a moderate correlation with measures of disability such as HAQ (Kendall's tau-b = 0.55,  $p < 0.001$ ). Unexpectedly, the correlation with the Medical Outcome Study Short-Form Survey 36 was moderate (Kendall's tau-b = -0.43,  $p = 0.001$ , for the mental component and Kendall's tau-b = -0.51,  $p = 0.001$ , for the physical component). Evidently the correlation with measures of activity (DAS28, tender and swollen joint count, self-report of pain, and global assessments) was poor (0.17 to 0.41).

The DRP showed a moderate correlation with measures

of disability such as the HAQ (Spearman's correlation = 0.59,  $p < 0.001$ ) and measures of handicap such as the LHS (Spearman's correlation = -0.51,  $p < 0.001$ ). The LHS showed good correlation with disability as measured by the HAQ (-0.71,  $p < 0.001$ ; the correlation is negative because the questionnaires are ordered in different directions). The correlation with disease activity measures was low (-0.36 to 0.01).

The HAQ was tested only against the LHS (Spearman's correlation = 0.71,  $p < 0.001$ ) and the DRP (Spearman's correlation = 0.59,  $p < 0.001$ ) as a similar construct, showing moderate to good correlations. The HAQ in UPIA does not have a good correlation with disease activity measures (0.17–0.41), although they are larger than for the DRP or the LHS. In summary, construct validity of these questionnaires is not bad for what they intend to measure, but they clearly do not measure disease activity.

With regard to the McROMI and NOAR-DJC, they are presented as instruments that should measure a construct close to disease activity. The problem with UPIA is that the number of swollen joints varies much more than in a specific disease (oligo to polyarthritis versus polyarthritis in RA, for instance), and relying on the number of swollen joints for disease activity may not be completely adequate. When the McROMI was tested against measures of disease activity, the correlation was poor, being the higher one when it was compared with DAS28-C-reactive protein (tau-b = 0.42,  $p < 0.001$ ). Correlation was not better when compared to measures of function, which best correlated with the HAQ (tau-b = 0.44,  $p < 0.001$ ). The NOAR-DJC was tested against different activity measures after 1 and 5 years of followup. Results after 1 year of followup showed low correlations: tender joint count,  $r = 0.18$  (95% CI 0.12, 0.24); swollen joint count,  $r = 0.21$  (95% CI 0.16, 0.27); HAQ,  $r = 0.39$  (95% CI 0.34, 0.44); and eroded joint count,  $r = 0.19$  (95% CI 0.10, 0.27). After 5-year followup, correlations were slightly better: tender joint count,  $r = 0.28$  (95% CI 0.20, 0.35); swollen joint count,  $r = 0.33$  (95% CI 0.25, 0.39); HAQ,  $r = 0.45$  (95% CI 0.40, 0.50); and eroded joint count,  $r = 0.42$  (95% CI 0.35, 0.49). This article actually shows the correlation in a subgroup that developed RA after 1 year: they were the same or worse.

The RADAI was tested only in UPIA versus the DAS28 (Pearson's correlation 0.596,  $p < 0.0001$ ). This was the best correlation with a disease activity measure that we found.

## DISCUSSION

Despite their well established use to evaluate disease activity in RA, it remains unclear whether composite indices may also be useful in patients with UPIA.

Our systematic review summarized available evidence from the literature on the most suitable clinical instruments to evaluate the diagnosis and followup of UPIA. Our results showed that no disease activity instrument has been fully



validated for use in UPIA; lack of validation is particularly apparent for the most commonly used indexes such as DAS28, SDAI, and CDAI, which should be included in a research agenda. We observed that since the questionnaires retrieved are designed mainly to measure disability, although they were in part validated in a UPIA population, they cannot be recommended to evaluate disease activity, not to mention diagnostic evolution and followup in these patients. The indices may be useful to evaluate disability in patients with UPIA and indirectly to evaluate disease progression. Physical disability is the most powerful determinant of all severe longterm outcomes in RA<sup>19</sup>, and possibly in UPIA as well.

Concerning the indices retrieved, only the RADAI, a mixed questionnaire-index, seems to be useful. However, it was not completely validated. Construct validity was examined versus the DAS28 only, showing a good correlation. This finding suggests that the RADAI may be a valid and feasible instrument for the assessment of disease activity in patients with UPIA, although it would be necessary to establish whether it can detect clinically important changes.

The 2 other indices are clearly not suitable, as they are time-consuming and difficult to perform, and validation is clearly incomplete. We are unaware of the effect size, and the construct validity is not very promising.

Validation of an instrument is a continuing process, and testing validity is established not from a single approach but from a series of converging studies. Future validation of these indices in different populations is necessary, especially because validity of an instrument is population-specific<sup>20</sup>.

There are important limitations to our analyses; the study of validity of any clinical measure in UPIA is very challenging, in particular since "disease activity" implies that a defined disease should be diagnosed, which is not the case in UPIA. Further, many of the included studies do not evaluate all aspects of the validation for an instrument.

In this systematic review we have not found direct evidence on what the most useful index is to follow up patients with UPIA, or at which intervals these should be repeated. The experts decided, in light of very limited evidence and mainly based on their experience, that disease activity should be monitored; however, no specific instrument can be recommended.

Future research is needed to evaluate the validity of the most common indices in populations with UPIA. Also, it would be important to adjust these indices to the different characteristics of these patients, possibly including larger joint counts or other extraarticular features.

## ACKNOWLEDGMENT

We thank all members of the 3e scientific committees, all participants of the national meetings, and acknowledge the support from Margaux Orange.

## REFERENCES

1. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van

- Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. *Arthritis Rheum* 2005;52:3381-90.
2. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364:263-9.
3. Machado P, Castrejon I, Katchamart W, Koevoets R, Kuriya B, Schoels M, et al. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3e Initiative. *Ann Rheum Dis* 2010 August 19. [Epub ahead of print]
4. Visser K, Katchamart W, Loza E, Martinez-Lopez JA, Salliot C, Trudeau J, et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3e Initiative. *Ann Rheum Dis* 2009;68:1086-93.
5. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine (Phila Pa 1976)* 2003;28:1290-9.
6. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71-2.
7. De Bruin AF, Diederiks JP, de Witte LP, Stevens FC, Philipsen H. Assessing the responsiveness of a functional status measure: the Sickness Impact Profile versus the SIP68. *J Clin Epidemiol* 1997;50:529-40.
8. Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49:916-20.
9. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* 2003;42:244-57.
10. Baron M, Schieir O, Hudson M, Steele R, Kolahi S, Berkson L, et al. The clinimetric properties of the World Health Organization Disability Assessment Schedule II in early inflammatory arthritis. *Arthritis Rheum* 2008;59:382-90.
11. Harwood RH, Rogers A, Dickinson E, Ebrahim S. Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. *Qual Health Care* 1994;3:11-6.
12. Carr AJ, Thompson PW. Towards a measure of patient-perceived handicap in rheumatoid arthritis. *Br J Rheumatol* 1994;33:378-82.
13. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789-93.
14. Mason JH, Anderson JJ, Meenan RF, Haralson KM, Lewis-Stevens D, Kaine JL. The Rapid Assessment of Disease Activity in Rheumatology (RADAR) questionnaire. Validity and sensitivity to change of a patient self-report measure of joint count and clinical status. *Arthritis Rheum* 1992;35:156-62.
15. Bykerk VP, Mirnyuk L, Chen H, Maguire L, Jamal S, Bombardier C. Validity of the RADAI in early rheumatoid arthritis [abstract]. *Ann Rheum Dis* 2007;66 Suppl 2:344.
16. Baron M, Steele R. Development of the McGill Range of Motion Index. *Clin Orthop Relat Res* 2007;456:42-50.
17. Bunn DK, Shepstone L, Galpin LM, Wiles NJ, Symmons DP. The NOAR Damaged Joint Count (NOAR-DJC): a clinical measure for assessing articular damage in patients with early inflammatory polyarthritis including rheumatoid arthritis. *Rheumatology*

- 2004;43:1519-25.
18. Harwood RH, Carr AJ, Thompson PW, Ebrahim S. Handicap in inflammatory arthritis. *Br J Rheumatol* 1996;35:891-7.
19. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. *Ann Intern Med* 1994;120:26-34.
20. Pedhauzer E, Schmelkin L. Measurement, design and analysis: an integrated approach. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991:37-9.
21. Tully M, Cantrill J. The test-retest reliability of the modified Patient Generated Index. *J Health Serv Res Policy* 2002;7:81-9.
22. Cohen JD, Dougados M, Goupille P, Cantagrel A, Meyer O, Sibilia J, et al. Health Assessment Questionnaire score is the best predictor of 5-year quality of life in early rheumatoid arthritis. *J Rheumatol* 2006;33:1936-41.
23. El Miedany Y, Palmer D. Can standard rheumatology clinical practice be patient-based? *Br J Nurs* 2008;17:673-5.
24. Kievit W, Welsing PM, Adang EM, Eijssbouts AM, Krabbe PF, van Riel PL. Comment on the use of self-reporting instruments to assess patients with rheumatoid arthritis: the longitudinal association between the DAS28 and the VAS general health. *Arthritis Rheum* 2006;55:745-50.
25. Suurmeijer TP, Doeglas DM, Moum T, Briancon S, Krol B, Sanderman R, et al. The Groningen Activity Restriction Scale for measuring disability: its utility in international comparisons. *Am J Public Health* 1994;84:1270-3.
26. Salaffi F, Stancati A, Neri R, Grassi W, Bombardieri S. Measuring functional disability in early rheumatoid arthritis: the validity, reliability and responsiveness of the Recent-Onset Arthritis Disability (ROAD) index. *Clin Exp Rheumatol* 2005;23:S31-42.
27. Lerner D, Amick BC 3rd, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. *Med Care* 2001;39:72-85.
28. Smolen JS, Aletaha D. Activity assessments in rheumatoid arthritis. *Curr Opin Rheumatol* 2008;20:306-13.
29. Hamilton J, Brydson G, Fraser S, Grant M. Walking ability as a measure of treatment effect in early rheumatoid arthritis. *Clin Rehabil* 2001;15:142-7.
30. Cole JC, Khanna D, Clements PJ, Seibold JR, Tashkin DP, Paulus HE, et al. Single-factor scoring validation for the Health Assessment Questionnaire-Disability Index (HAQ-DI) in patients with systemic sclerosis and comparison with early rheumatoid arthritis patients. *Qual Life Res* 2006;15:1383-94.
31. Saraux A, Berthelot JM, Chales G, Le Henaff C, Thorel JB, Hoang S, et al. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. *Arthritis Rheum* 2001;44:2485-91.
32. Van der Helm-van Mil AH, Detert J, le Cessie S, Filer A, Bastian H, Burmester GR, et al. Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. *Arthritis Rheum* 2008;58:2241-7.
33. Van der Helm-van Mil AH, le Cessie S, van Dongen H, Breedveld FC, Toes RE, Huizinga TW. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;56:433-40.
34. El Miedany Y, Youssef S, Mehanna AN, El Gaafary M. Development of a scoring system for assessment of outcome of early undifferentiated inflammatory synovitis. *Joint Bone Spine* 2008;75:155-62.

**2.2. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. (Ann Rheum Dis 2011;70:15-24)**





# Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative

P Machado, I Castrejon, W Katchamart, et al.

*Ann Rheum Dis* published online August 19, 2010  
doi: 10.1136/ard.2010.130625

---

Updated information and services can be found at:  
<http://ard.bmj.com/content/early/2010/08/17/ard.2010.130625.full.html>

---

*These include:*

- |                               |  |
|-------------------------------|--|
| <b>References</b>             | This article cites 173 articles, 68 of which can be accessed free at:<br><a href="http://ard.bmj.com/content/early/2010/08/17/ard.2010.130625.full.html#ref-list-1">http://ard.bmj.com/content/early/2010/08/17/ard.2010.130625.full.html#ref-list-1</a> |
| <b>Open Access</b>            | This paper is freely available online under the BMJ Journals unlocked scheme, see <a href="http://ard.bmj.com/info/unlocked.dtl">http://ard.bmj.com/info/unlocked.dtl</a>  |
| <b>P&lt;P</b>                 | Published online August 19, 2010 in advance of the print journal.  |
| <b>Email alerting service</b> | Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.   |
- 

**Topic collections** Articles on similar topics can be found in the following collections  
[Unlocked](#) (459 articles)

---

## Notes

---

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To order reprints of this article go to:  
<http://ard.bmj.com/cgi/reprintform>

To subscribe to *Annals of the Rheumatic Diseases* go to:  
<http://ard.bmj.com/subscriptions>

# Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative

P Machado,<sup>1,2</sup> I Castrejon,<sup>3</sup> W Katchamart,<sup>4</sup> R Koevoets,<sup>2</sup> B Kuriya,<sup>4,5</sup> M Schoels,<sup>6</sup> L Silva-Fernández,<sup>7</sup> K Thevissen,<sup>8</sup> W Vercourte,<sup>9</sup> E Villeneuve,<sup>10</sup> D Aletaha,<sup>11</sup> L Carmona,<sup>12</sup> R Landewé,<sup>13</sup> D van der Heijde,<sup>2</sup> J W J Bijlsma,<sup>14</sup> V Bykerk,<sup>4</sup> H Canhão,<sup>15</sup> A I Catrina,<sup>16</sup> P Durez,<sup>17</sup> C J Edwards,<sup>18</sup> M D Mjaavatten,<sup>19</sup> B F Leeb,<sup>20</sup> B Losada,<sup>21</sup> E M Martín-Mola,<sup>22</sup> P Martínez-Osuna,<sup>23</sup> C Montecucco,<sup>24</sup> U Müller-Ladner,<sup>25</sup> M Østergaard,<sup>26</sup> B Sheane,<sup>27</sup> R M Xavier,<sup>28</sup> J Zochling,<sup>29</sup> C Bombardier<sup>30,31</sup>

For numbered affiliations see end of article

## Correspondence to

Dr Pedro Machado,  
Department of Rheumatology,  
Coimbra University Hospital,  
Praceta Mota Pinto,  
3000-075 Coimbra, Portugal;  
[pedromcmachado@gmail.com](mailto:pedromcmachado@gmail.com)

PM, IC, WK, RK, BK, MS, LS-F, KT, WV and EV contributed equally to this work.

Accepted 10 July 2010

## ABSTRACT

**Objective** To develop evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis (UPIA).

**Methods** 697 rheumatologists from 17 countries participated in the 3E (Evidence, Expertise, Exchange) Initiative of 2008–9 consisting of three separate rounds of discussions and modified Delphi votes. In the first round 10 clinical questions were selected. A bibliographic team systematically searched Medline, Embase, the Cochrane Library and ACR/EULAR 2007–2008 meeting abstracts. Relevant articles were reviewed for quality assessment, data extraction and synthesis. In the second round each country elaborated a set of national recommendations. Finally, multinational recommendations were formulated and agreement among the participants and the potential impact on their clinical practice was assessed.

**Results** A total of 39 756 references were identified, of which 250 were systematically reviewed. Ten multinational key recommendations about the investigation and follow-up of UPIA were formulated. One recommendation addressed differential diagnosis and investigations prior to establishing the operational diagnosis of UPIA, seven recommendations related to the diagnostic and prognostic value of clinical and laboratory assessments in established UPIA (history and physical examination, acute phase reactants, autoantibodies, radiographs, MRI and ultrasound, genetic markers and synovial biopsy), one recommendation highlighted predictors of persistence (chronicity) and the final recommendation addressed monitoring of clinical disease activity in UPIA.

**Conclusions** Ten recommendations on how to investigate and follow-up UPIA in the clinical setting were developed. They are evidence-based and supported by a large panel of rheumatologists, thus enhancing their validity and practical use.

## INTRODUCTION

In clinical practice, a large number of patients who present with recent-onset arthritis have

undifferentiated peripheral inflammatory arthritis (UPIA). In this context, patients' initial questions will focus on their likelihood of developing a well-defined rheumatic disease and on what the future holds for disease progression, persistence, functional impairment and quality of life. These are questions about future diagnosis and prognosis. The answers to these questions are vital for clinical decision making, including the choice of treatment.

The 3E Initiative (Evidence, Expertise, Exchange) in rheumatology is a multinational effort aimed at promoting evidence-based medicine by formulating practical recommendations addressing clinical problems.<sup>1 2</sup> The objective of the 3E Initiative of 2008–9 was to develop practical recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis by integrating systematically generated evidence and expert opinion of a broad panel of international rheumatologists. Although the term 'inflammatory' in UPIA may seem redundant, the reason for its use was to clearly distinguish the target population from patients with degenerative joint disease, often called osteoarthritis or degenerative arthritis in the English medical literature.

## METHODS

A total of 697 rheumatologists from 17 countries participated in the 3E Initiative of 2008–9. Each country was represented by a scientific committee consisting of one principal investigator and 5–13 members. The bibliographic team consisted of 10 international fellows (PM, IC, WK, RK, BK, MS, LS-F, KT, WV, EV) and five mentors (DA, LC, RL, DvdH, CB), one of the mentors also being the scientific organiser (CB). The 17 national principal investigators were selected and invited by the 3E scientific organiser (CB) and each national chair was in charge of composing a national steering committee. The experts were all the members of



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

## Recommendations

the 17 national steering committees who attended the multinational meetings for the 3E Initiative.

During the first international meeting (n=113 participants), 10 clinically relevant questions on how to investigate and follow-up UPIA were formulated and selected via a modified Delphi vote. The areas addressed were fourfold: (1) the phase prior to establishing the operational diagnosis of UPIA—namely, which differential diagnosis should be considered in a patient presenting with (inflammatory) arthritis and the minimal investigations necessary to consider a patient as having UPIA; (2) the diagnostic and prognostic value of clinical assessment and investigations in UPIA (history and physical examination, acute phase reactants, autoantibodies, x-rays, MRI, ultrasound (US), genetic markers and synovial biopsy); (3) the predictors of persistence (chronicity) in UPIA; and (4) the measures of clinical disease activity in UPIA.

The clinical questions were structured using the PIO format (Patients, Participants or Problem; Intervention or Index test; Outcomes or target conditions).<sup>3</sup> The patients included 'adults with UPIA'. Duration of symptoms was not an exclusion criterion. The definition of UPIA is controversial and there is no widely accepted classification criterion for this condition. During the 2008–9 3E Initiative kick-off meeting, experts decided that only patients in whom clinically apparent joint swelling (synovial proliferation or synovial effusion) was observed by the rheumatologist should be included. For our review we systematically searched for studies of patients who did not fulfil diagnostic/classification criteria for any specific rheumatic disorder after initial assessment. Studies with mixed populations (eg, UPIA+arthralgia, UPIA+early rheumatoid arthritis (RA)) were also retained, as these could be useful for extrapolating results. The intervention or index test was defined according to each question (eg, erosions on x-rays, anti-citrullinated protein/peptide antibodies (ACPA) positivity) and the index test should have been assessed at baseline. The outcomes were defined as the development of well-defined rheumatic diseases (eg, RA, psoriatic arthritis) or relevant disease outcomes (eg, remission, radiographic progression). As diagnostic/classification criteria we accepted either internationally validated criteria (eg, American College of Rheumatology criteria for RA<sup>4</sup>) or the opinion of the treating physician/investigator.

A systematic literature search for articles published up to February 2009 was carried out in Medline, Embase and Cochrane Library using comprehensive search strategies elaborated in collaboration with experienced librarians. The searches were limited to diagnostic and prognostic studies using a modification

of published sensitive search strategies.<sup>5–8</sup> No language restrictions were used. Retrieved citations were screened for titles, abstracts and full text using predefined inclusion and exclusion criteria; full read papers and review articles were hand-searched for additional references. Retained articles were graded for their methodological quality according to the levels of evidence of the Oxford Centre for Evidence-Based Medicine (<http://www.cebm.net/index.aspx?o=1025>).

Each question was addressed separately by independent searches. For each question, relevant data were extracted and appropriate statistics were calculated, including OR, sensitivity, specificity, positive/negative predictive values and positive/negative likelihood ratios. Details and results of the literature search for each question will be published separately, while the current article describes the merging process between the evidence found for each question and the interpretation of this by the experts, having the 10 recommendations as the result.

In the second round, a national meeting was held in each country (total=697 participants) to discuss the generated evidence and propose a set of recommendations. In a third joint meeting the 17 scientific committees (n=94 participants) merged all propositions into 10 final recommendations via discussion and modified Delphi vote. The grade of recommendation according to the Oxford levels of evidence was attributed and the level of agreement was measured on a 10-point numerical rating scale (1=no agreement, 10=full agreement).<sup>9</sup> Finally, the potential effect of each recommendation in clinical practice was assessed according to three impact statements voted by the rheumatologists.

## RESULTS

A total of 39 756 references were identified, of which 250 were systematically reviewed (table 1). The 10 multinational key recommendations are listed in table 2 with the corresponding level of evidence and grade of recommendation. The mean level of agreement among the rheumatologists was 8.7 (range 7.4–9.1). The percentage of rheumatologists who indicated they would change their clinical practice according to each recommendation is shown in table 3. Evidence for repeating investigations was not found for any of the questions, therefore all recommendations about this topic were based on expert opinion.

**Recommendation 1.** *All possible causes of arthritis (idiopathic, autoimmune, degenerative, infectious, malignancy, traumatic, metabolic) should be considered in the differential diagnosis. Complete history and thorough physical examination will determine the ranking order of possible differential diagnoses. Investigations should be based on the differential diagnosis of the patient.*

**Table 1** Results of the systematic literature search for each recommendation topic

Recommendation (number and topic)	Retrieved references by systematic literature search (n)	Articles included in the systematic reviews (n)
1. Pre-UIA differential diagnosis and investigations	540	51
2. History and physical examination	2914	37
3. Acute phase reactants	3699	18
4. Autoantibodies	13217	64
5. X-rays	3585	25
6.1. MRI	2595	11
6.2. Ultrasound	2111	2
7. Genetic markers	3109	26
8. Synovial biopsy	6536	4
9. Predictors of persistence (chronicity)	437	7
10. Measures of clinical disease activity	1013	5
Total	39756	250

UPIA, undifferentiated peripheral inflammatory arthritis.

**Table 2** Multinational recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis

Recommendation (with level of evidence and grade of recommendation)	Agreement mean (SD)
1. All possible causes of arthritis (idiopathic, autoimmune, degenerative, infectious, malignancy, traumatic, metabolic) should be considered in the differential diagnosis. Complete history and thorough physical examination will determine the ranking order of possible differential diagnoses [5, D]. Investigations should be based on the differential diagnosis of the patient [5, D]	9.0 (1.7)
2. To establish a specific diagnosis and prognosis following presentation of UPIA, a careful systematic history and physical examination should be performed, with particular attention to age, gender [1a, A], geographical area [5, D], functional status [1a, A], duration of symptoms/early morning stiffness, number plus pattern of tender/swollen joints [1a, A], axial/enthesal involvement and extra-articular/systemic features [5, D]	8.8 (1.3)
3. ESR and CRP should be performed at baseline in the investigation for diagnosis [2b, B] and prognosis [2b, B] of UPIA and repeated when clinically relevant [5, D]	9.1 (1.4)
4. Testing of RF and/or ACPA should be performed in the evaluation of patients with UPIA, as these factors are predictive of RA diagnosis and prognosis; negative tests do not exclude progression to RA [1a, A]. If a connective tissue disease/systemic inflammatory disorder is suspected, additional autoantibody tests should be considered [5, D]	9.1 (1.2)
5. X-rays of affected joints should be performed at baseline [5, D]. X-rays of hands, wrists and feet should be considered in the evaluation of UPIA as the presence of erosions is predictive for the development of RA and persistence of disease [1a, A]. These should be repeated within 1 year [5, D]	7.4 (2.6)
6. There is insufficient evidence to recommend the routine use of MRI and US for diagnosis or prognosis in UPIA [5, D]; in UPIA and suspicion of RA, MRI of hands and wrists could be considered for diagnosis [2b, B]	8.2 (2.0)
7. There is no genetic test that can be routinely recommended [3b, D], however HLA-B27 testing may be helpful in specific clinical settings [5, D]	8.8 (1.5)
8. Routine synovial biopsy is not recommended but can give information for differential diagnosis, especially in patients with persistent monoarthritis [2b, B]	8.8 (1.8)
9. Predictors of persistent inflammatory arthritis should be documented and include disease duration of $\geq 6$ weeks [1b, A], morning stiffness $> 30$ min [4, C], functional impairment [4, C], involvement of small joints [4, C] and/or knee [4, C], involvement of $\geq 3$ joints [1b, B], ACPA [4, C] and/or RF positivity [4, C] and presence of radiographic erosion [1b, B]	8.6 (1.7)
10. Disease activity should be monitored [5, D], however no specific tool can be recommended [3b, C]	9.0 (1.7)

Values in square brackets indicate [level of evidence, grade of recommendation] according to the Oxford Centre for Evidence-based Medicine levels of evidence.

Agreement was voted on a scale from 1 to 10 (fully disagree to fully agree) by the 94 rheumatologists attending the 3E Multi-National Closing Meeting. These attendees were members of the 17 scientific committees involved in the 3E Initiative of 2008–2009.

ACPA, anti-citrullinated protein/peptide antibodies; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; RA, rheumatoid arthritis; RF, rheumatoid factor; UPIA, undifferentiated peripheral inflammatory arthritis; US, ultrasound.

**Table 3** Percentage of rheumatologists in the 3E Initiative who indicated for each recommendation if it would change their clinical practice

Recommendation (number and topic)	The recommendation will change my practice (%)	The recommendation is already my practice (%)	I don't want to change my practice for this aspect (%)
1. Pre-UIPA differential diagnosis and investigations	0	96.5	3.5
2. History and physical examination	0	98.3	1.8
3. Acute phase reactants	5.4	91.1	3.6
4. Autoantibodies	1.8	96.4	1.8
5. X-rays	16.1	48.2	35.7
6. MRI and ultrasound	17.9	64.3	17.9
7. Genetic markers	1.8	92.9	5.4
8. Synovial biopsy	8.9	83.9	7.1
9. Predictors of persistence (chronicity)	24.6	66.7	8.8
10. Measures of clinical disease activity	12.3	84.2	3.5

UIPA, undifferentiated peripheral inflammatory arthritis.

As UPIA is an operational diagnosis after excluding well-defined rheumatic diseases, the question about pre-UIPA differential diagnosis and investigations was analysed by looking at the diagnosis that was excluded in cohorts of patients with UPIA and by identifying the inclusion and exclusion criteria of these studies as well as the investigations performed before the UPIA cohort was established. RA was the most frequent diagnosis reported as exclusion criterion<sup>10–59</sup> and there was no standard baseline investigation undertaken prior to inclusion as UPIA (table 4).<sup>41–60</sup>

Experts agreed that, when facing a new patient presenting with arthritis, every diagnosis needed to be kept in mind as UPIA is an exclusion diagnosis. Although the consensus was that it was impossible to name all possible diagnoses, it was felt useful to mention some major disease categories to make sure that these are considered. Experts also advised that UPIA should be constantly rethought, as patients may develop a disease that can be labelled with a specific diagnosis at any time. Moreover, this recommendation applies only if arthritis persists and not if it is self-limiting. Again, as the investigations will vary according to context and clinical presentation, experts felt that it would not be useful to make a list of recommended minimal investigations.

Recommendation 2. *To establish a specific diagnosis and prognosis following presentation of UPIA, a careful systematic history and physical examination should be performed with particular attention to age, gender, geographical area, functional status, duration of symptoms/early morning stiffness, number plus pattern of tender/swollen joints, axial/enthesal involvement and extra-articular/systemic features.*

Although selected observational studies were of good quality, there was large heterogeneity with respect to the type of history and physical examination features described.<sup>39 40 42–49 61–87</sup> Of the quantified features, advanced age,<sup>44 83</sup> female gender<sup>44</sup> and greater morning stiffness<sup>43 44</sup> were predictive of an eventual diagnosis of RA. A higher number of tender<sup>44</sup> and swollen joints,<sup>43 44 61</sup> involvement of small joints of hands and feet,<sup>44 83</sup> involvement of both the upper and lower extremities<sup>44</sup> and symmetrical involvement<sup>43</sup> were also associated with progression to RA. Similar features were associated with disease persistence<sup>81–87</sup> and development of erosions,<sup>48 63 78</sup> while self-reported functional disability (Health Assessment Questionnaire (HAQ) score)<sup>67 76</sup> and the presence of extra-articular features<sup>76</sup> were uniquely predictive of future disability, along with advanced age,<sup>67 76</sup> female gender<sup>67</sup> and longer symptom duration.<sup>67</sup>



## Recommendations

**Table 4** Diagnosis reported as exclusion criteria and baseline investigations undertaken prior to inclusion as UPIA (ordered by the frequency of reporting in the retrieved literature), both in studies including patients exclusively with UPIA as well as in selected mixed populations that included a well-defined subset of patients with UPIA

<b>A. Reported differential diagnosis prior to establishing the operational diagnosis of UPIA</b>	
- Rheumatoid arthritis	- Polymyalgia rheumatica
- Osteoarthritis	- Lyme disease
- Spondyloarthritis (reactive arthritis, psoriatic arthritis, ankylosing spondylitis and undifferentiated spondyloarthritis)	- Vasculitis
- Crystal-related arthritis	- Juvenile inflammatory arthritis
- Trauma	- Palindromic rheumatism
- Connective tissue diseases (systemic lupus erythematosus, Sjögren syndrome and myositis)	- Fibromyalgia
- Septic arthritis	- Endocrinological origin
- Sarcoidosis	- Malignancy-related arthritis
- Soft tissue disorders	- Viral aetiology
<b>B. Reported investigations prior to establishing the operational diagnosis of UPIA</b>	
- History	- Microbiological assessment
- Tender and swollen joint count	- Anti-citrullinated protein/peptide antibodies
- Rheumatoid factor	- Radiography of the chest and/or of other affected joints
- C reactive protein	- Urinalysis
- Physical examination	- Thyroid function tests
- Hands and feet x-rays	- C3, C4
- Full blood count	- Immunoglobulins
- Antinuclear antibodies	- Antibodies to extractable nuclear antigens
- Erythrocyte sedimentation rate	- Antibodies to double-stranded DNA
- Biochemistry (liver function tests, glucose, urate and renal function)	- Specific serological assessment
- HLA typing (HLA-B27 and HLA-DR)	

UPIA, undifferentiated peripheral inflammatory arthritis.

Experts recognised the importance of the abovementioned evidence-based features and, based on their clinical experience, also highlighted the contribution of the patient's geographical area of residence, the presence of axial/enthesal involvement and the presence of extra-articular/systemic features. However, the greater relevance given to features included in the recommendation does not preclude the need to perform a careful systematic history and physical examination in every patient with UPIA.

**Recommendation 3.** *Erythrocyte sedimentation rate (ESR) and C reactive protein (CRP) should be performed at baseline in the investigation for diagnosis and prognosis of UPIA and repeated when clinically relevant.*

Elevated erythrocyte sedimentation rate (ESR) showed some diagnostic value for the development of RA<sup>74 85</sup> but no prognostic value for persistence (chronicity) or structural damage.<sup>40 45 88</sup> C reactive protein (CRP) appeared to be a poor predictor of persistent arthritis, radiological progression and functional disability.<sup>80 89</sup> However, there was some evidence for the usefulness of elevated CRP in predicting RA, especially when the CRP levels are higher.<sup>48 88</sup> In one study, CRP did not have any diagnostic value with regard to spondylarthropathy.<sup>39</sup> For other acute phase reactants, the evidence on diagnostic or prognostic value was scarce, negative or controversial.<sup>32 42 48 79 80 90–95</sup>

Based on sparse evidence and on personal experience regarding acute phase reactants, experts recommended that only ESR and CRP should be performed at baseline and repeated according to the clinical setting.

**Recommendation 4.** *Testing of rheumatoid factor (RF) and/or ACPA should be performed in the evaluation of patients with UPIA, as these factors are predictive of RA diagnosis and prognosis; negative tests do not exclude progression to RA. If a connective tissue disease/systemic inflammatory disorder is suspected, additional autoantibody tests should be considered.*

The association of ACPA and rheumatoid factor (RF)<sup>11 42–44 48 50 73 96–110</sup> with a diagnosis of RA at follow-up was compelling in the retrieved literature. The absence of ACPA or RF was diagnostically less helpful. The presence of ACPA or RF<sup>75 106–109 111–115</sup>

also increased the probability of developing persistent synovitis or a worse radiographic outcome.<sup>73 75 84–86 116</sup> For anti-keratin antibodies (AKA) and anti-perinuclear factor, the evidence suggests diagnostic usefulness; AKA also appears to have some prognostic value.<sup>11 96–99 107 110 114 117</sup> For all other markers including a variety of other autoantibodies as well as bone and cartilage biomarkers, the evidence for diagnostic or prognostic value is scarce, negative or controversial.<sup>57 102 118–126</sup> The same applies to disease outcomes different from those already mentioned.<sup>59 74 76 81 93 100 116 127 128</sup>

The value of ACPA and RF in UPIA was recognised and, based on clinical experience, experts also advised consideration of additional autoantibody tests if non-RA systemic inflammatory disorders are suspected. The use of the general term ACPA was preferred as the literature describes several tests for detecting antibodies to citrullinated peptides (such as anti-CCP1 and anti-CCP2) and newer generation tests are also expected to be used in the future.

**Recommendation 5.** *X-rays of affected joints should be performed at baseline. X-rays of hands, wrists and feet should be considered in the evaluation of UPIA as the presence of erosions is predictive for the development of RA and persistence of disease. These should be repeated within 1 year.*

Radiographic erosions<sup>43 49</sup> and Larsen grade 1 (in a population without erosions at baseline)<sup>20</sup> increased the probability of developing RA from UPIA. Moreover, when comparing mild versus progressive disease after 1 year follow-up, Sharp/van der Heijde scores at baseline were significantly higher in the progressive disease group.<sup>48</sup> In another study,<sup>44</sup> erosions were found to be a predictor of RA in univariate but not in multivariate analysis.

Overall, studies in mixed populations also provided some evidence for the usefulness of x-rays in predicting RA.<sup>72 88 92 109 122 129–135</sup> In general, prognosis was worse when radiographic abnormalities at baseline were more severe.<sup>75 91 109 116 133 136–140</sup>

Experts recognised the clinical value of hand and feet x-rays in UPIA and, based on clinical experience, also recommended that x-rays of affected joints should be performed at baseline;

furthermore, experts advised that x-rays should be repeated within 1 year (in case of disease persistence). Moreover, although not voted to be included in the recommendation, some of the experts expressed their opinion that pelvic/sacroiliac joint x-rays should also be considered, particularly in RF- and ACPA-negative patients or if spondyloarthritis is suspected.

There was a slightly lower agreement about this recommendation (table 2, 7.4 agreement), with a larger proportion of experts stating that they did not want to change their practice for this aspect (table 3, 35.7%). This lower concordance was mainly related to the inclusion of 'x-rays of affected joints at baseline' and about the advice to repeat x-rays 'within 1 year'.

**Recommendation 6.** *There is insufficient evidence to recommend the routine use of magnetic resonance imaging (MRI) and ultrasound (US) for diagnosis or prognosis in UPIA; in UPIA and suspicion of RA, MRI of hands and wrists could be considered for diagnosis.*

Bone oedema was found to be an independent predictor of the future development of RA from UPIA,<sup>141</sup> and the presence of a distinct MRI synovitis and erosion pattern with the involvement of several hand joints but not the first carpometacarpal joint also increased the probability of developing RA.<sup>20</sup> The absence of the same MRI synovitis pattern decreased the probability of developing RA.<sup>20</sup> Overall, MRI studies in mixed populations<sup>101 134 142–147</sup> provided some evidence for the usefulness of MRI (bone oedema, synovitis and erosions) in predicting RA. Regarding US, two mixed populations revealed US-power Doppler signal and US-gray scale synovitis as potential candidates for future studies in UPIA.<sup>148 149</sup>

Experts recognised that MRI of the hands and wrists has already been shown to be useful in predicting the development of RA from UPIA, while the value of US in UPIA is still to be determined. However, data are still too scarce to recommend the routine use of any of these imaging tools. This recommendation does not dispute the fact that, compared with physical examination and x-rays, both MRI and US may offer advantages through more sensitive depiction of inflammatory and destructive disease manifestations. The current recommendation pertains only to the diagnostic and prognostic value of these imaging tools in UPIA.

**Recommendation 7.** *There is no genetic test that can be routinely recommended, however HLA-B27 testing may be helpful in specific clinical settings.*

There was a great heterogeneity among the genetic markers tested.<sup>39 40 46 50–52 65 84 127 133 150–165</sup> The shared epitope (SE) was the most frequently studied marker. Eight studies<sup>40 50 65 133 153–155 158</sup> tested its diagnostic utility and showed poor results. Only in one study was the positive likelihood ratio for RA relevant, but this result came from the study with the poorest quality and smallest sample size.<sup>40</sup> In isolation, no other genetic marker was informative of a future diagnosis in patients with UPIA. With regard to prognosis, the SE was weakly associated with a poor prognosis of arthritis in terms of development of erosions, mortality, disability and persistent synovitis.<sup>65 127 133 163 164</sup> Other genes were not good predictors of erosions or other less studied outcomes.

The experts acknowledged the current lack of evidence for the practical utility of genetics in UPIA. However, based on their clinical experience, experts chose to highlight that HLA-B27 may be helpful in the appropriate clinical setting—namely, when spondyloarthritis is suspected.

**Recommendation 8.** *Routine synovial biopsy is not recommended but can give information for differential diagnosis, especially in patients with persistent monoarthritis.*

Studies had significant clinical and statistical heterogeneity.<sup>22 23 166 167</sup> Three broad synovial features of interest were identified in the literature: ACPA staining, immunohistochemistry and vascular patterns. In contrast to serological ACPA testing, ACPA staining was shown not to be highly specific for a diagnosis of RA.<sup>167</sup> In one study, synovial histopathology seemed to differentiate between RA and non-RA.<sup>166</sup> The vascular pattern in undifferentiated arthritis was not specific enough to differentiate between spondyloarthritis and RA.<sup>22 23</sup>

The exact role of synovial biopsy in UPIA is yet to be determined and experts felt that it could not be recommended as a routine procedure. However, experts also highlighted the fact that synovial biopsy may give important diagnostic clues, especially in some selected cases (eg, persistent/chronic refractory monoarthritis, suspicion of malignancy or suspicion of chronic infection such as tuberculosis).

**Recommendation 9.** *Predictors of persistent inflammatory arthritis should be documented and include disease duration of ≥6 weeks, morning stiffness >30 min, functional impairment, involvement of small joints and/or knee, involvement of ≥3 joints, ACPA and/or RF positivity and presence of radiographic erosion.*

The question about chronicity was investigated by looking at prognostic studies that used multivariate analysis to identify independent predictors of persistence (chronicity). At baseline the following variables were found to be independent predictors of persistent (inflammatory) arthritis: disease duration,<sup>75 82 116</sup> duration of morning stiffness,<sup>75 85 86</sup> change of functional status (measured by HAQ) in the first 3 months,<sup>82</sup> failure to respond 2 weeks after local treatment with intra-articular corticosteroids,<sup>82</sup> small joint involvement,<sup>168</sup> knee involvement,<sup>85</sup> presence of RF,<sup>75 85</sup> presence and level of ACPA,<sup>75 86 168</sup> functional status (HAQ),<sup>169</sup> arthritis of at least three joints,<sup>75</sup> proximal interphalangeal joint involvement,<sup>169</sup> metatarsophalangeal joint involvement<sup>75</sup> and radiographic erosion at the hands and feet.<sup>75</sup> The magnitude of the association in the same predictor was diverse among the studies depending on the patient characteristics (namely, if the population was purely UPIA or not), the study design and the variables used to adjust for in the models.

**Recommendation 10.** *Disease activity should be monitored, however no specific tool can be recommended.*

Five studies evaluated the validation of different clinical measures in patients with UPIA. Validation aspects of four questionnaires (WHO Disability Assessment Schedule,<sup>170</sup> London Handicap Scale, Disease Repercussion Profile and the HAQ<sup>171</sup>) and three physical measures (RA Disease Activity Index,<sup>172</sup> McGill Range of Motion Index<sup>173</sup> and NOAR Damage Joint Count<sup>174</sup>) were partially assessed in these studies but none of the instruments of disease activity was fully validated for its use in UPIA.

Although no instrument of disease activity has been fully validated for its use in UPIA, experts felt that it was important to recommend that there should be a conscious effort to record disease activity.

## DISCUSSION

Ten multinational recommendations on how to investigate and follow-up UPIA in the clinical setting were developed, which are practical, evidence-based and supported by a large panel of international rheumatologists in the 3E Initiative.

We followed an established group decision method. A representative expert panel of 697 academic and community rheumatologists from 17 countries selected relevant questions that reflect the challenges of approaching a patient with UPIA. They

## Recommendations

openly discussed the evidence from the literature followed by a silent voting process. We used the touch pad methodology with prespecified cut-off levels of agreement to generate the final recommendations. Several rounds of rewording and revoting were sometimes required to reach the specified cut-off for agreement. This process highlights the international dimension of this collaboration and strengthens the current recommendations.<sup>1 2</sup> It ensured that the final recommendations were evidence-driven as well as clinically relevant.

Furthermore, the broad participation increases external validity and enhances future dissemination and implementation into rheumatological practice worldwide. Another main feature of the 3E Initiative was the promotion of epidemiology and systematic literature research, all participants having been updated on how to appraise published evidence.

There is widespread interest in predictive medicine. Following a strict methodology, we aimed to find all available evidence regarding each question which resulted in a large number of reviewed articles. However, the evidence in truly UPIA populations is scarce, exposing the need to create a research agenda addressing this topic. In particular, future studies should clearly distinguish between individuals with early well-defined rheumatic diseases, individuals with UPIA and individuals with inflammatory joint symptoms but no obvious joint swelling. All these populations can be studied for predictive algorithms and results may be different depending on the study population.

The definition of UPIA is controversial and much of the literature is skewed towards early RA. The difficulty in defining UPIA is underlined by the continuous changing face of different categories of patients, which can be well illustrated by the recent new ACR/EULAR criteria for RA,<sup>175</sup> as several of the patients we now describe as having UPIA will likely be labelled as having RA. Nevertheless, despite the influence that this changing may have on research and daily practice, the recommendations presented in this article are based on currently available evidence. They may help the clinician in the effective management of patients with UPIA and can be adjusted if future studies or clinical experience reveal new insights.

In summary, multinational recommendations for the investigation and follow-up of patients with undifferentiated arthritis in daily clinical practice were developed, integrating systematic literature review and expert opinion with the aim of promoting evidence-based medicine and ultimately improving patient care.

**Acknowledgements** The authors thank all members of the 3E scientific committees, all participants of the national meetings, the support from Margaux Orange and the librarians who helped in elaborating the systematic literature searches. CB holds a Canada Research Chair in Knowledge Transfer for Musculoskeletal Care.

**Funding** This work was supported by Abbott with an unrestricted educational grant.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Author affiliations** <sup>1</sup>Rheumatology Department, Coimbra University Hospital, Coimbra, Portugal

<sup>2</sup>Rheumatology Department, Leiden University Medical Center, Leiden, The Netherlands

<sup>3</sup>Rheumatology Department, Hospital Universitario de La Princesa, Madrid, Spain

<sup>4</sup>Mount Sinai Hospital, Toronto, Ontario, Canada

<sup>5</sup>Harvard School of Public Health, Boston, Massachusetts, USA

<sup>6</sup>KH Hietzing, Vienna, Austria

<sup>7</sup>Rheumatology Department, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain

<sup>8</sup>Rheumatology Department, Maastricht University Medical Center, Maastricht, The Netherlands

<sup>9</sup>Rheumatology Department, Atrium Medical Centre Parkstad, Heerlen, The Netherlands

<sup>10</sup>Academic Unit of Musculoskeletal Disease, Leeds Institute of Molecular Medicine, Leeds University, Leeds, UK

<sup>11</sup>Medical University, Vienna, Austria

<sup>12</sup>Research Unit, Fundación Española de Reumatología, Madrid, Spain

<sup>13</sup>Department of Internal Medicine/Rheumatology, Maastricht University Medical Center, Care and Public Health Research Institute, Maastricht, The Netherlands

<sup>14</sup>Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>15</sup>Rheumatology Research Unit, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Serviço de Reumatologia e Doenças Ósseas Metabólicas, Hospital de Santa Maria, Lisbon, Portugal

<sup>16</sup>Department of Rheumatology, Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden

<sup>17</sup>Microbiology Department, Cliniques universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium

<sup>18</sup>Southampton University Hospital, Southampton, UK

<sup>19</sup>Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

<sup>20</sup>State Hospital Stockerau, Center for Rheumatology, Lower Austria, Stockerau, Austria

<sup>21</sup>Hospital Universitario de Caracas, Ciudad Universitaria, Los Chaguaramos, Caracas, Venezuela

<sup>22</sup>Hospital Universitario La Paz, Department of Rheumatology, Universidad Autónoma, Madrid, Spain

<sup>23</sup>Rheumatology Unit, Centro de Investigación Farmacológica y Biotecnológica, Hospital y Fundación Clínica Médica Sur, México City, México

<sup>24</sup>Cattedra di Reumatologia, IRCCS Policlinico S. Matteo, Università di Pavia, Pavia, Italy

<sup>25</sup>Department of Rheumatology and Clinical Immunology, Justus-Liebig-University Gießen, Kerckhoff Clinic, Bad Nauheim, Germany

<sup>26</sup>Department of Rheumatology, Copenhagen University Hospitals at Hvidovre and Glostrup, Hvidovre, Denmark

<sup>27</sup>Department of Rheumatology, St James's Hospital, Dublin, Ireland

<sup>28</sup>Hospital de Clínicas de Porto Alegre, Faculdade de Medicina da Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

<sup>29</sup>Menzies Research Institute, University of Tasmania, Hobart, Australia

<sup>30</sup>Division of Rheumatology, University Health Network, Toronto, Ontario, Canada

<sup>31</sup>Department of Health Policy, Management, and Evaluation, University of Toronto, Toronto, Ontario, Canada

## REFERENCES

1. **Sidiropoulos PI**, Hatemi G, Song IH, *et al*. Evidence-based recommendations for the management of ankylosing spondylitis: systematic literature search of the 3E Initiative in Rheumatology involving a broad panel of experts and practising rheumatologists. *Rheumatology (Oxford)* 2008;**47**:355–61.
2. **Visser K**, Katchamart W, Loza E, *et al*. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Ann Rheum Dis* 2009;**68**:1086–93.
3. **Sackett DL**, Richardson WS, Rosenberg WM, *et al*. *Evidence-based medicine: how to practice and teach EBM*. London, UK: Churchill Livingstone, 1997.
4. **Arnett FC**, Edworthy SM, Bloch DA, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.
5. **Haynes RB**, McKibbon KA, Wilczynski NL, *et al*. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;**330**:1179.
6. **Wilczynski NL**, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med* 2004;**2**:23.
7. **Wilczynski NL**, Haynes RB. EMBASE search strategies for identifying methodologically sound diagnostic studies for use by clinicians and researchers. *BMC Med* 2005;**3**:7.
8. **Wilczynski NL**, Haynes RB. Optimal search strategies for detecting clinically sound prognostic studies in EMBASE: an analytic survey. *J Am Med Inform Assoc* 2005;**12**:481–5.
9. **Roddy E**, Zhang W, Doherty M, *et al*. Evidence-based clinical guidelines: a new system to better determine true strength of recommendation. *J Eval Clin Pract* 2006;**12**:347–52.
10. **Savolainen E**, Kaipiainen-Seppänen O, Kröger L, *et al*. Total incidence and distribution of inflammatory joint diseases in a defined population: results from the Kuopio 2000 arthritis survey. *J Rheumatol* 2003;**30**:2460–8.
11. **Berthelot JM**, Maugars Y, Castagné A, *et al*. Antiperinuclear factors are present in polyarthritis before ACR criteria for rheumatoid arthritis are fulfilled. *Ann Rheum Dis* 1997;**56**:123–5.
12. **Blaauw I**, Dijkmans B, Bouma P, *et al*. Rational diagnosis and treatment in unclassified arthritis: how clinical data may guide requests for Lyme serology and antibiotic treatment. *Ann Rheum Dis* 1993;**52**:206–10.



13. **O'Hara R**, Murphy EP, Whitehead AS, *et al*. Local expression of the serum amyloid A and formyl peptide receptor-like 1 genes in synovial tissue is associated with matrix metalloproteinase production in patients with inflammatory arthritis. *Arthritis Rheum* 2004;**50**:1788–99.
14. **Parker JD**, Capell HA. An acute arthritis clinic—one year's experience. *Br J Rheumatol* 1986;**25**:293–5.
15. **Rooney T**, Murphy E, Benito M, *et al*. Synovial tissue interleukin-18 expression and the response to treatment in patients with inflammatory arthritis. *Ann Rheum Dis* 2004;**63**:1393–8.
16. **Emad Y**, Ragab Y, Shaarawy A, *et al*. Enhanced MRI in early undifferentiated oligoarthritis of the knee joints: improvements already visible after 2 months of DMARDs treatment. *Clin Rheumatol* 2008;**27**:1177–82.
17. **Appel H**, Mertz A, Distler A, *et al*. The 19 kDa protein of *Yersinia enterocolitica* O:3 is recognized on the cellular and humoral level by patients with *Yersinia* induced reactive arthritis. *J Rheumatol* 1999;**26**:1964–71.
18. **Inaoui R**, Bertin P, Preux PM, *et al*. Outcome of patients with undifferentiated chronic monoarthritis: retrospective study of 46 cases. *Joint Bone Spine* 2004;**71**:209–13.
19. **Kaarela K**, Tiitinen S, Luukkainen R. Long-term prognosis of monoarthritis. A follow-up study. *Scand J Rheumatol* 1983;**12**:374–6.
20. **Duer A**, Østergaard M, Hørslev-Petersen K, *et al*. Magnetic resonance imaging and bone scintigraphy in the differential diagnosis of unclassified arthritis. *Ann Rheum Dis* 2008;**67**:48–51.
21. **Schnarr S**, Putschky N, Jendro MC, *et al*. Chlamydia and *Borrelia* DNA in synovial fluid of patients with early undifferentiated oligoarthritis: results of a prospective study. *Arthritis Rheum* 2001;**44**:2679–85.
22. **Baeten D**, Kruithof E, De Rycke L, *et al*. Diagnostic classification of spondylarthropathy and rheumatoid arthritis by synovial histopathology: a prospective study in 154 consecutive patients. *Arthritis Rheum* 2004;**50**:2931–41.
23. **Cañete JD**, Rodríguez JR, Salvador G, *et al*. Diagnostic usefulness of synovial vascular morphology in chronic arthritis. A systematic survey of 100 cases. *Semin Arthritis Rheum* 2003;**32**:378–87.
24. **Pazdur J**, Ploski R, Bogunia-Kubik K, *et al*. Can HLA-DRB1 typing have prognostic value in patients with undifferentiated chronic arthritis? *Tissue Antigens* 1998;**51**:678–80.
25. **Higami K**, Hakoda M, Matsuda Y, *et al*. Lack of association of HLA-DRB1 genotype with radiologic progression in Japanese patients with early rheumatoid arthritis. *Arthritis Rheum* 1997;**40**:2241–7.
26. **Wilkinson NZ**, Kingsley GH, Sieper J, *et al*. Lack of correlation between the detection of Chlamydia trachomatis DNA in synovial fluid from patients with a range of rheumatic diseases and the presence of an antichlamydial immune response. *Arthritis Rheum* 1998;**41**:845–54.
27. **Zavala-Cerna MG**, Nava A, García-Castañeda E, *et al*. Serum IgG activity against cyclic citrullinated peptide in patients evaluated for rheumatoid factor correlates with the IgM isotype. *Rheumatol Int* 2008;**28**:851–7.
28. **Braun J**, Laitko S, Treharne J, *et al*. Chlamydia pneumoniae—a new causative agent of reactive arthritis and undifferentiated oligoarthritis. *Ann Rheum Dis* 1994;**53**:100–5.
29. **Braun J**, Tuszewski M, Ehlers S, *et al*. Nested polymerase chain reaction strategy simultaneously targeting DNA sequences of multiple bacterial species in inflammatory joint diseases. II. Examination of sacroiliac and knee joint biopsies of patients with spondyloarthropathies and other arthritides. *J Rheumatol* 1997;**24**:1101–5.
30. **Dryll A**, Lansaman J, Cazalis P, *et al*. Light and electron microscopy study of capillaries in normal and inflammatory human synovial membrane. *J Clin Pathol* 1977;**30**:556–62.
31. **Fendler C**, Laitko S, Sörensen H, *et al*. Frequency of triggering bacteria in patients with reactive arthritis and undifferentiated oligoarthritis and the relative importance of the tests used for diagnosis. *Ann Rheum Dis* 2001;**60**:337–43.
32. **Hitchon CA**, Alex P, Erdile LB, *et al*. A distinct multicytokine profile is associated with anti-cyclical citrullinated peptide antibodies in patients with early untreated inflammatory arthritis. *J Rheumatol* 2004;**31**:2336–46.
33. **Jendro MC**, Raum E, Schnarr S, *et al*. Cytokine profile in serum and synovial fluid of arthritis patients with Chlamydia trachomatis infection. *Rheumatol Int* 2005;**25**:37–41.
34. **Jones VE**, Jacoby RK, Cowley PJ, *et al*. Immune complexes in early arthritis. II. Immune complex constituents are synthesized in the synovium before rheumatoid factors. *Clin Exp Immunol* 1982;**49**:31–40.
35. **Jones V**, Taylor PC, Jacoby RK, *et al*. Synovial synthesis of rheumatoid factors and immune complex constituents in early arthritis. *Ann Rheum Dis* 1984;**43**:235–9.
36. **Shine B**, Bourne JT, Begum Baig F, *et al*. C reactive protein and immunoglobulin G in synovial fluid and serum in joint disease. *Ann Rheum Dis* 1991;**50**:32–5.
37. **Siala M**, Jaulhac B, Gdoura R, *et al*. Analysis of bacterial DNA in synovial tissue of Tunisian patients with reactive and undifferentiated arthritis by broad-range PCR, cloning and sequencing. *Arthritis Res Ther* 2008;**10**:R40.
38. **Nissilä M**, Isomäki H, Kaarela K, *et al*. Prognosis of inflammatory joint diseases. A three-year follow-up study. *Scand J Rheumatol* 1983;**12**:33–8.
39. **Kvien TK**, Glennäs A, Melby K. Prediction of diagnosis in acute and subacute oligoarthritis of unknown origin. *Br J Rheumatol* 1996;**35**:359–63.
40. **Morel J**, Legouffe MC, Bozonat MC, *et al*. Outcomes in patients with incipient undifferentiated arthritis. *Joint Bone Spine* 2000;**67**:49–53.
41. **Verpoort KN**, Jol-van der Zijde CM, Papendrecht-van der Voort EA, *et al*. Isotype distribution of anti-cyclic citrullinated peptide antibodies in undifferentiated arthritis and rheumatoid arthritis reflects an ongoing immune response. *Arthritis Rheum* 2006;**54**:3799–808.
42. **Savolainen E**, Kautiainen H, Koivula MK, *et al*. Change of diagnoses and outcome of patients with early inflammatory joint diseases during a mean 13-month follow-up. *Scand J Rheumatol* 2007;**36**:194–7.
43. **van Gaalen FA**, Linn-Rasker SP, van Venrooij WJ, *et al*. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. *Arthritis Rheum* 2004;**50**:709–15.
44. **van der Helm-van Mil AH**, le Cessie S, van Dongen H, *et al*. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;**56**:433–40.
45. **Zeidler H**, Werdier D, Klauder A, *et al*. Undifferentiated arthritis and spondylarthropathy as a challenge for prospective follow-up. *Clin Rheumatol* 1987;**6**(Suppl 2):112–20.
46. **Hülsemann JL**, Zeidler H. Undifferentiated arthritis in an early synovitis out-patient clinic. *Clin Exp Rheumatol* 1995;**13**:37–43.
47. **Machold KP**, Stamm TA, Eberl GJ, *et al*. Very recent onset arthritis—clinical, laboratory, and radiological findings during the first year of disease. *J Rheumatol* 2002;**29**:2278–87.
48. **Jansen LM**, van Schaardenburg D, van der Horst-Bruinsma IE, *et al*. One year outcome of undifferentiated polyarthritis. *Ann Rheum Dis* 2002;**61**:700–3.
49. **van Aken J**, van Dongen H, le Cessie S, *et al*. Comparison of long term outcome of patients with rheumatoid arthritis presenting with undifferentiated arthritis or with rheumatoid arthritis: an observational cohort study. *Ann Rheum Dis* 2006;**65**:20–5.
50. **van der Helm-van Mil AH**, Verpoort KN, Breedveld FC, *et al*. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. *Arthritis Rheum* 2006;**54**:1117–21.
51. **Feitsma AL**, Toes RE, Begovich AB, *et al*. Risk of progression from undifferentiated arthritis to rheumatoid arthritis: the effect of the PTPN22 1858T-allele in anti-citrullinated peptide antibody positive patients. *Rheumatology (Oxford)* 2007;**46**:1092–5.
52. **Wesoly J**, Hu X, Thabet MM, *et al*. The 620W allele is the PTPN22 genetic variant conferring susceptibility to RA in a Dutch population. *Rheumatology (Oxford)* 2007;**46**:617–21.
53. **Stahl HD**, Seidl B, Hubner B, *et al*. High incidence of parvovirus B19 DNA in synovial tissue of patients with undifferentiated mono- and oligoarthritis. *Clin Rheumatol* 2000;**19**:281–6.
54. **van Dongen H**, van Aken J, Lard LR, *et al*. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum* 2007;**56**:1424–32.
55. **van der Helm-van Mil AH**, Verpoort KN, le Cessie S, *et al*. The HLA-DRB1 shared epitope alleles differ in the interaction with smoking and predisposition to antibodies to cyclic citrullinated peptide. *Arthritis Rheum* 2007;**56**:425–32.
56. **Verpoort KN**, van Gaalen FA, van der Helm-van Mil AH, *et al*. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. *Arthritis Rheum* 2005;**52**:3058–62.
57. **Matsumoto I**, Lee DM, Goldbach-Mansky R, *et al*. Low prevalence of antibodies to glucose-6-phosphate isomerase in patients with rheumatoid arthritis and a spectrum of other chronic autoimmune disorders. *Arthritis Rheum* 2003;**48**:944–54.
58. **Wilbrink B**, van der Heijden IM, Schouls LM, *et al*. Detection of bacterial DNA in joint samples from patients with undifferentiated arthritis and reactive arthritis, using polymerase chain reaction with universal 16S ribosomal RNA primers. *Arthritis Rheum* 1998;**41**:535–43.
59. **Visser K**, Verpoort KN, van Dongen H, *et al*. Pretreatment serum levels of anti-cyclic citrullinated peptide antibodies are associated with the response to methotrexate in recent-onset arthritis. *Ann Rheum Dis* 2008;**67**:1194–5.
60. **Saleem B**, Mackie S, Quinn M, *et al*. Does the use of tumour necrosis factor antagonist therapy in poor prognosis, undifferentiated arthritis prevent progression to rheumatoid arthritis? *Ann Rheum Dis* 2008;**67**:1178–80.
61. **Alarcón GS**, Willkens RF, Ward JR, *et al*. Early undifferentiated connective tissue disease. IV. Musculoskeletal manifestations in a large cohort of patients with undifferentiated connective tissue diseases compared with cohorts of patients with well-established connective tissue diseases: followup analyses in patients with unexplained polyarthritis and patients with rheumatoid arthritis at baseline. *Arthritis Rheum* 1996;**39**:403–14.
62. **Binard A**, Alassane S, Devauchelle-Pensec V, *et al*. Outcome of early monoarthritis: a followup study. *J Rheumatol* 2007;**34**:2351–7.
63. **Bukhari M**, Lunt M, Barton A, *et al*. Increasing age at symptom onset is associated with worse radiological damage at presentation in patients with early inflammatory polyarthritis. *Ann Rheum Dis* 2007;**66**:389–93.
64. **Devlin J**, Gough A, Huissoon A, *et al*. The outcome of knee synovitis in early arthritis provides guidelines for management. *Clin Rheumatol* 2000;**19**:82–5.



## Recommendations

65. **El-Gabalawy HS**, Goldbach-Mansky R, Smith D 2nd, *et al.* Association of HLA alleles and clinical features in patients with synovitis of recent onset. *Arthritis Rheum* 1999;**42**:1696–705.
66. **Gerber LH**, Furst G, Yarbrough C, *et al.* Number of active joints, not diagnosis, is the primary determinant of function and performance in early synovitis. *Clin Exp Rheumatol* 2003;**21**(5 Suppl 31):S65–70.
67. **Glennäs A**, Kvien TK, Andrup O, *et al.* Recent onset arthritis in the elderly: a 5 year longitudinal observational study. *J Rheumatol* 2000;**27**:101–8.
68. **Harrison BJ**, Symmons DP, Brennan P, *et al.* Inflammatory polyarthritis in the community is not a benign disease: predicting functional disability one year after presentation. *J Rheumatol* 1996;**23**:1326–31.
69. **Hernandez-Avila M**, Liang MH, Willett WC, *et al.* Exogenous sex hormones and the risk of rheumatoid arthritis. *Arthritis Rheum* 1990;**33**:947–53.
70. **Hernández Avila M**, Liang MH, Willett WC, *et al.* Reproductive factors, smoking, and the risk for rheumatoid arthritis. *Epidemiology* 1990;**1**:285–91.
71. **Jensen T**, Klarlund M, Hansen M, *et al.* Bone loss in unclassified polyarthritis and early rheumatoid arthritis is better detected by digital x ray radiogrammetry than dual x ray absorptiometry: relationship with disease activity and radiographic outcome. *Ann Rheum Dis* 2004;**63**:15–22.
72. **Kuriya B**, Cheng CK, Chen HM, *et al.* Validation of a prediction rule for development of rheumatoid arthritis in patients with early undifferentiated arthritis. *Ann Rheum Dis* 2009;**68**:1482–5.
73. **Quinn MA**, Green MJ, Marzo-Ortega H, *et al.* Prognostic factors in a large cohort of patients with early undifferentiated inflammatory arthritis after application of a structured management protocol. *Arthritis Rheum* 2003;**48**:3039–45.
74. **Schumacher HR**, Jr, Habre W, Meador R, *et al.* Predictive factors in early arthritis: long-term follow-up. *Semin Arthritis Rheum* 2004;**33**:264–72.
75. **Visser H**, le Cessie S, Vos K, *et al.* How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002;**46**:357–65.
76. **Wiles NJ**, Dunn G, Barrett EM, *et al.* One year followup variables predict disability 5 years after presentation with inflammatory polyarthritis with greater accuracy than at baseline. *J Rheumatol* 2000;**27**:2360–6.
77. **Luchikhina EL**, Karateev DE, Nasonov EL. Recent onset of inflammatory arthritis in different age groups. *Ann Rheum Dis* 2007;**66**(Suppl 2):331.
78. **van der Horst-Bruinsma IE**, Speyer I, Visser H, *et al.* Diagnosis and course of early-onset arthritis: results of a special early arthritis clinic compared to routine patient care. *Br J Rheumatol* 1998;**37**:1084–8.
79. **Wolfe F**, Ross K, Hawley DJ, *et al.* The prognosis of rheumatoid arthritis and undifferentiated polyarthritis syndrome in the clinic: a study of 1141 patients. *J Rheumatol* 1993;**20**:2005–9.
80. **Woolf AD**, Hall ND, Goulding NJ, *et al.* Predictors of the long-term outcome of early synovitis: a 5-year follow-up study. *Br J Rheumatol* 1991;**30**:251–4.
81. **Green M**, Marzo-Ortega H, McGonagle D, *et al.* Persistence of mild, early inflammatory arthritis: the importance of disease duration, rheumatoid factor, and the shared epitope. *Arthritis Rheum* 1999;**42**:2184–8.
82. **Green M**, Marzo-Ortega H, Wakefield RJ, *et al.* Predictors of outcome in patients with oligoarthritis: results of a protocol of intraarticular corticosteroids to all clinically active joints. *Arthritis Rheum* 2001;**44**:1177–83.
83. **Mjaavatten MD**, Nygaard H, Haugen AJ, *et al.* Baseline predictors of persistent arthritis, DMARD start and rheumatoid arthritis diagnosis: one year follow-up of 395 patients with very early arthritis. *Arthritis Rheum* 2008;**58**:1633.
84. **Stockman A**, Tait BD, Wolfe R, *et al.* Clinical, laboratory and genetic markers associated with erosions and remission in patients with early inflammatory arthritis: a prospective cohort study. *Rheumatol Int* 2006;**26**:500–9.
85. **Tunn EJ**, Bacon PA. Differentiating persistent from self-limiting symmetrical synovitis in an early arthritis clinic. *Br J Rheumatol* 1993;**32**:97–103.
86. **El Miedany Y**, Youssef S, Mehanna AN, *et al.* Development of a scoring system for assessment of outcome of early undifferentiated inflammatory synovitis. *Joint Bone Spine* 2008;**75**:155–62.
87. **Harrison BJ**, Symmons DP, Brennan P, *et al.* Natural remission in inflammatory polyarthritis: issues of definition and prediction. *Br J Rheumatol* 1996;**35**:1096–100.
88. **Jensen T**, Klarlund M, Hansen M, *et al.* Connective tissue metabolism in patients with unclassified polyarthritis and early rheumatoid arthritis. Relationship to disease activity, bone mineral density, and radiographic outcome. *J Rheumatol* 2004;**31**:1698–708.
89. **Mjaavatten MD**, Nygaard H, Haugen AJ, *et al.* Disease characteristics and predictors of persistent arthritis after one year in a very early arthritis clinic in Norway. *Ann Rheum Dis* 2007;**66**(Suppl II):332.
90. **Jansen LM**, van der Horst-Bruinsma IE, van Schaardenburg D, *et al.* Comparison of the baseline disease activity of early oligo- and polyarthritis in sequential years. *Clin Exp Rheumatol* 2004;**22**:447–52.
91. **Jansen LM**, van der Horst-Bruinsma I, Lems WF, *et al.* Serological bone markers and joint damage in early polyarthritis. *J Rheumatol* 2004;**31**:1491–6.
92. **Knudsen LS**, Klarlund M, Skjødt H, *et al.* Biomarkers of inflammation in patients with unclassified polyarthritis and early rheumatoid arthritis. Relationship to disease activity and radiographic outcome. *J Rheumatol* 2008;**35**:1277–87.
93. **Reneses S**, Pestana L, Fernandez-Suarez A, *et al.* A recent onset inflammatory polyarthritis register in Spain: factors that predict remission. *Scand J Rheumatol* 2007;**36**:378–85.
94. **Hall ND**, Blake DR, Bacon PA. Serum sulphhydryl levels in early synovitis. *J Rheumatol* 1982;**9**:593–6.
95. **Alexander GJ**, Blake DR, Holman RL, *et al.* Predictive value of paired plasma and serum viscosity in early rheumatic conditions. *BMJ (Clin Res Ed)* 1981;**282**:1198.
96. **Saraux A**, Berthelot JM, Chalès G, *et al.* Value of laboratory tests in early prediction of rheumatoid arthritis. *Arthritis Rheum* 2002;**47**:155–65.
97. **Devauchelle-Pensec V**, Saraux A, Youinou P, *et al.* Antiperinuclear factor and antikeratin/antifilaggrin antibodies for differentiating early rheumatoid arthritis from polymyalgia rheumatica. *Joint Bone Spine* 2001;**68**:306–10.
98. **Goldbach-Mansky R**, Lee J, McCoy A, *et al.* Rheumatoid arthritis associated autoantibodies in patients with synovitis of recent onset. *Arthritis Res* 2000;**2**:236–43.
99. **Cordonnier C**, Meyer O, Palazzo E, *et al.* Diagnostic value of anti-RA33 antibody, antikeratin antibody, antiperinuclear factor and antinuclear antibody in early rheumatoid arthritis: comparison with rheumatoid factor. *Br J Rheumatol* 1996;**35**:620–4.
100. **Verstappen SMM**, McCoy MJ, Roberts C, *et al.* Predictors of poor prognosis in very early inflammatory polyarthritis. *Arthritis Rheum* 2008;**58**:S769.
101. **Tamai M**, Kawakami A, Uetani M, *et al.* Early prediction of rheumatoid arthritis by serological variables and magnetic resonance imaging of the wrists and finger joints: results from prospective clinical examination. *Ann Rheum Dis* 2006;**65**:134–5.
102. **van der Helm-van Mil AH**, Detert J, le Cessie S, *et al.* Towards personalized medicine in rheumatology – a prediction rule for the development of rheumatoid arthritis in patients with undifferentiated arthritis. *Arthritis Rheum* 2008;**58**:S917.
103. **Raza K**, Breese M, Nightingale P, *et al.* Predictive value of antibodies to cyclic citrullinated peptide in patients with very early inflammatory arthritis. *J Rheumatol* 2005;**32**:231–8.
104. **Jansen AL**, van der Horst-Bruinsma I, van Schaardenburg D, *et al.* Rheumatoid factor and antibodies to cyclic citrullinated Peptide differentiate rheumatoid arthritis from undifferentiated polyarthritis in patients with early arthritis. *J Rheumatol* 2002;**29**:2074–6.
105. **van der Helm-van Mil AH**, Detert J, le Cessie S, *et al.* Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. *Arthritis Rheum* 2008;**58**:2241–7.
106. **Nell VP**, Machold KP, Stamm TA, *et al.* Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis. *Ann Rheum Dis* 2005;**64**:1731–6.
107. **Aho K**, Palosuo T, Lukka M, *et al.* Antifilaggrin antibodies in recent-onset arthritis. *Scand J Rheumatol* 1999;**28**:113–16.
108. **Schellekens GA**, Visser H, de Jong BA, *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum* 2000;**43**:155–63.
109. **Nielen MM**, van der Horst AR, van Schaardenburg D, *et al.* Antibodies to citrullinated human fibrinogen (ACF) have diagnostic and prognostic value in early arthritis. *Ann Rheum Dis* 2005;**64**:1199–204.
110. **Vittecoq O**, Jouen-Beades F, Krzanowska K, *et al.* Rheumatoid factors, anti-flaggrin antibodies and low in vitro interleukin-2 and interferon-gamma production are useful immunological markers for early diagnosis of community cases of rheumatoid arthritis. A preliminary study. *Joint Bone Spine* 2001;**68**:144–53.
111. **Hitchon CA**, Wong K, El-Gabalawy HS. Measurement of baseline serum matrix metalloproteinase levels adds minimal prognostic value over routine clinical parameters in the prediction of radiographic erosions in early inflammatory arthritis. *Arthritis Rheum* 2008;**58**:S754.
112. **Fèvre C**, Brazier M, Daragon A, *et al.* Can we predict structural damage progression at 2 years in very early arthritis? Value of bone and cartilage markers in the conservatively treated community-based inception VERA cohort. *Ann Rheum Dis* 2007;**66**(Suppl II):324.
113. **Bukhari M**, Thomson WV, Naseem H, *et al.* The performance of anti-cyclic citrullinated peptide antibodies in predicting the severity of radiologic damage in inflammatory polyarthritis: results from the Norfolk Arthritis Register. *Arthritis Rheum* 2007;**56**:2929–35.
114. **Kurki P**, von Essen R, Kaarela K, *et al.* Antibody to stratum corneum (antikeratin antibody) and antiperinuclear factor: markers for progressive rheumatoid arthritis. *Scand J Rheumatol* 1997;**26**:346–9.
115. **Teitsson I**, Withington RH, Seifert MH, *et al.* Prospective study of early rheumatoid arthritis. I. Prognostic value of IgA rheumatoid factor. *Ann Rheum Dis* 1984;**43**:673–8.
116. **Boire G**, Cossette P, de Brum-Fernandes AJ, *et al.* Anti-Sa antibodies and antibodies against cyclic citrullinated peptide are not equivalent as predictors of severe outcomes in patients with recent-onset polyarthritis. *Arthritis Res Ther* 2005;**7**:R592–603.
117. **Vittecoq O**, Incauragarat B, Jouen-Beades F, *et al.* Autoantibodies recognizing citrullinated rat flaggrin in an ELISA using citrullinated and non-citrullinated recombinant proteins as antigens are highly diagnostic for rheumatoid arthritis. *Clin Exp Immunol* 2004;**135**:173–80.

118. **Jouen F**, Vittecoq O, Leguillou F, *et al*. Diagnostic and prognostic values of anti glucose-6-phosphate isomerase antibodies in community-recruited patients with very early arthritis. *Clin Exp Immunol* 2004;**137**:606–11.
119. **Vittecoq O**, Salle V, Jouen-Beades F, *et al*. Autoantibodies to the 27 C-terminal amino acids of calpastatin are detected in a restricted set of connective tissue diseases and may be useful for diagnosis of rheumatoid arthritis in community cases of very early arthritis. *Rheumatology (Oxford)* 2001;**40**:1126–34.
120. **Goldbach-Mansky R**, Lee JM, Hoxworth JM, *et al*. Active synovial matrix metalloproteinase-2 is associated with radiographic erosions in patients with early synovitis. *Arthritis Res* 2000;**2**:145–53.
121. **Goldbach-Mansky R**, Suson S, Wesley R, *et al*. Raised granzyme B levels are associated with erosions in patients with early rheumatoid factor positive rheumatoid arthritis. *Ann Rheum Dis* 2005;**64**:715–21.
122. **Cunhane G**, Fitzgerald O, Beeton C, *et al*. Early joint erosions and serum levels of matrix metalloproteinase 1, matrix metalloproteinase 3, and tissue inhibitor of metalloproteinases 1 in rheumatoid arthritis. *Arthritis Rheum* 2001;**44**:2263–74.
123. **Kudo-Tanaka E**, Ohshima S, Ishii M, *et al*. Autoantibodies to cyclic citrullinated peptide 2 (CCP2) are superior to other potential diagnostic biomarkers for predicting rheumatoid arthritis in early undifferentiated arthritis. *Clin Rheumatol* 2007;**26**:1627–33.
124. **Boire G**, Abrahamowicz M, King LE, *et al*. Association between serum biomarkers of cartilage turnover and radiographic and symptomatic progression in an early polyarticular inflammatory arthritis cohort. *Ann Rheum Dis* 2007;**66**(Suppl II):322.
125. **Patel S**, Farragher T, Berry J, *et al*. Association between serum vitamin D metabolite levels and disease activity in patients with early inflammatory polyarthritis. *Arthritis Rheum* 2007;**56**:2143–9.
126. **Berthelot JM**, Saraux A, Audrain M, *et al*. Poor predictive value of antinucleosome and antineutrophil cytoplasmic antibodies in a 270 inception cohort of patients with early naked arthritis of less than one year's duration. *Ann Rheum Dis* 2002;**61**:760–1.
127. **Farragher TM**, Goodson NJ, Naseem H, *et al*. Association of the HLA-DRB1 gene with premature death, particularly from cardiovascular disease, in patients with rheumatoid arthritis and inflammatory polyarthritis. *Arthritis Rheum* 2008;**58**:359–69.
128. **Ortiz AM**, González-Álvarez I, García-Vicuña R, *et al*. Anti-cyclic citrullinated peptide antibodies and high IL-15 serum levels predict better than rheumatoid factor the requirement of intensive treatment in early arthritis patients. *Ann Rheum Dis* 2007;**66**(Suppl II):593.
129. **Devauchelle Pensec V**, Saraux A, Berthelot JM, *et al*. Ability of hand radiographs to predict a further diagnosis of rheumatoid arthritis in patients with early arthritis. *J Rheumatol* 2001;**28**:2603–7.
130. **Devauchelle Pensec V**, Saraux A, Berthelot JM, *et al*. Ability of foot radiographs to predict rheumatoid arthritis in patients with early arthritis. *J Rheumatol* 2004;**31**:66–70.
131. **Saraux A**, Berthelot JM, Chalès G, *et al*. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. *Arthritis Rheum* 2001;**44**:2485–91.
132. **Devauchelle-Pensec V**, Berthelot JM, Jousse S, *et al*. Performance of hand radiographs in predicting the diagnosis in patients with early arthritis. *J Rheumatol* 2006;**33**:1511–15.
133. **Gough A**, Faint J, Salmon M, *et al*. Genetic typing of patients with inflammatory arthritis at presentation can be used to predict outcome. *Arthritis Rheum* 1994;**37**:1166–70.
134. **Klarlund M**, Ostergaard M, Jensen KE, *et al*. Magnetic resonance imaging, radiography, and scintigraphy of the finger joints: one year follow up of patients with early arthritis. The TIRA Group. *Ann Rheum Dis* 2000;**59**:521–8.
135. **Daragon A**, Krzanowska K, Vittecoq O, *et al*. Prospective X-ray densitometry and ultrasonography study of the hand bones of patients with rheumatoid arthritis of recent onset. *Joint Bone Spine* 2001;**68**:34–42.
136. **Jansen LM**, van Schaardenburg D, van der Horst-Bruinsma I, *et al*. The predictive value of anti-cyclic citrullinated peptide antibodies in early arthritis. *J Rheumatol* 2003;**30**:1691–5.
137. **Bukhari M**, Lunt M, Harrison BJ, *et al*. Rheumatoid factor is the major predictor of increasing severity of radiographic erosions in rheumatoid arthritis: results from the Norfolk Arthritis Register Study, a large inception cohort. *Arthritis Rheum* 2002;**46**:906–12.
138. **Bukhari MA**, Wiles NJ, Lunt M, *et al*. Influence of disease-modifying therapy on radiographic outcome in inflammatory polyarthritis at five years: results from a large observational inception study. *Arthritis Rheum* 2003;**48**:46–53.
139. **Isomäki H**, Martio J, Sarna S, *et al*. Predicting the outcome of rheumatoid arthritis. A Soviet-Finnish co-operative study. *Scand J Rheumatol* 1984;**13**:33–8.
140. **Isomäki HA**. An epidemiologically based follow-up study of recent arthritis. Incidence, outcome and classification. *Clin Rheumatol* 1987;**6**(Suppl 2):53–9.
141. **Tamai M**, Kawakami A, Uetani M, *et al*. Anti-cyclic citrullinated peptide antibody and magnetic resonance imaging-detection of bone marrow oedema are most important predictors in classification as well as prognostic evaluation of undifferentiated arthritis. *Ann Rheum Dis* 2007;**66**(Suppl II):338.
142. **Mori G**, Tokunaga D, Takahashi KA, *et al*. Maximum intensity projection as a tool to diagnose early rheumatoid arthritis. *Mod Rheumatol* 2008;**18**:247–51.
143. **Narváez J**, Sirvent E, Narváez JA, *et al*. Usefulness of magnetic resonance imaging of the hand versus anticyclic citrullinated peptide antibody testing to confirm the diagnosis of clinically suspected early rheumatoid arthritis in the absence of rheumatoid factor and radiographic erosions. *Semin Arthritis Rheum* 2008;**38**:101–9.
144. **Solau-Gervais E**, Legrand JL, Cortet B, *et al*. Magnetic resonance imaging of the hand for the diagnosis of rheumatoid arthritis in the absence of anti-cyclic citrullinated peptide antibodies: a prospective study. *J Rheumatol* 2006;**33**:1760–5.
145. **Boutry N**, Hachulla E, Flipo RM, *et al*. MR imaging findings in hands in early rheumatoid arthritis: comparison with those in systemic lupus erythematosus and primary Sjögren syndrome. *Radiology* 2005;**236**:593–600.
146. **Sugimoto H**, Takeda A, Hyodoh K. Early-stage rheumatoid arthritis: prospective study of the effectiveness of MR imaging for diagnosis. *Radiology* 2000;**216**:569–75.
147. **Sugimoto H**, Takeda A, Masuyama J, *et al*. Early-stage rheumatoid arthritis: diagnostic accuracy of MR imaging. *Radiology* 1996;**198**:185–92.
148. **Freeston J**, Wakefield R, Conaghan P, *et al*. Ultrasound at presentation predicts clinical outcome in very early inflammatory patients. *Arthritis Rheum* 2007;**56**.
149. **Scire C**, Montecucco C, Epis O, *et al*. Residual disease activity assessment by musculoskeletal ultrasounds in early arthritis. *Arthritis Rheum* 2008;**58**:S408.
150. **Barton A**, Bowes J, Eyre S, *et al*. Investigation of polymorphisms in the PADI4 gene in determining severity of inflammatory polyarthritis. *Ann Rheum Dis* 2005;**64**:1311–15.
151. **Dubost JJ**, Demarquilly F, Soubrier M, *et al*. HLA and self-limiting, unclassified rheumatism. A role for HLA-B35? *J Rheumatol* 1999;**26**:2400–3.
152. **Saudan-Kister A**, Gabay C, Tiercy JM, *et al*. Adult seronegative arthritis with antinuclear antibodies: a distinct group of patients with a different immunogenetic pattern from seropositive rheumatoid arthritis and a good outcome. *Rev Rheum Engl Ed* 1996;**63**:313–20.
153. **Goëb V**, Dieudé P, Daveau R, *et al*. Contribution of PTPN22 1858T, TNFR1F 196R and HLA-shared epitope alleles with rheumatoid factor and anti-citrullinated protein antibodies to very early rheumatoid arthritis diagnosis. *Rheumatology (Oxford)* 2008;**47**:1208–12.
154. **Thomson W**, Harrison B, Ollier B, *et al*. Quantifying the exact role of HLA-DRB1 alleles in susceptibility to inflammatory polyarthritis: results from a large, population-based study. *Arthritis Rheum* 1999;**42**:757–62.
155. **Vos K**, van der Horst-Bruinsma IE, Hazes JM, *et al*. Evidence for a protective role of the human leukocyte antigen class II region in early rheumatoid arthritis. *Rheumatology (Oxford)* 2001;**40**:133–9.
156. **Willis G**, Scott DG, Jennings BA, *et al*. HFE mutations in an inflammatory arthritis population. *Rheumatology (Oxford)* 2002;**41**:176–9.
157. **Barton A**, Lamb R, Symmons D, *et al*. Macrophage migration inhibitory factor (MIF) gene polymorphism is associated with susceptibility to but not severity of inflammatory polyarthritis. *Genes Immun* 2003;**4**:487–91.
158. **Jacobsen S**, Madsen HO, Klarlund M, *et al*. The influence of mannose binding lectin polymorphisms on disease outcome in early polyarthritis. TIRA Group. *J Rheumatol* 2001;**28**:935–42.
159. **Nasrallah NS**, Masi AT, Chandler RW, *et al*. HLA-B27 antigen and rheumatoid factor negative (seronegative) peripheral arthritis. Studies in younger patients with early-diagnosed arthritis. *Am J Med* 1977;**63**:379–86.
160. **Naseem H**, Thomson W, Silman A, *et al*. The PTPN22\*1858T functional polymorphism is associated with susceptibility to inflammatory polyarthritis but neither this nor other variants spanning the gene is associated with disease outcome. *Ann Rheum Dis* 2008;**67**:251–5.
161. **Barton A**, Platt H, Salway F, *et al*. Polymorphisms in the tumour necrosis factor gene are not associated with severity of inflammatory polyarthritis. *Ann Rheum Dis* 2004;**63**:280–4.
162. **Barton A**, Platt H, Salway F, *et al*. Polymorphisms in the mannose binding lectin (MBL) gene are not associated with radiographic erosions in rheumatoid or inflammatory polyarthritis. *J Rheumatol* 2004;**31**:442–7.
163. **Emery P**, Salmon M, Bradley H, *et al*. Genetically determined factors as predictors of radiological change in patients with early symmetrical arthritis. *BMJ* 1992;**305**:1387–9.
164. **Harrison B**, Thomson W, Symmons D, *et al*. The influence of HLA-DRB1 alleles and rheumatoid factor on disease outcome in an inception cohort of patients with early inflammatory arthritis. *Arthritis Rheum* 1999;**42**:2174–83.
165. **John S**, Smith S, Morrison JF, *et al*. Genetic variation in CCR5 does not predict clinical outcome in inflammatory arthritis. *Arthritis Rheum* 2003;**48**:3615–16.
166. **Kraan MC**, Haringman JJ, Post WJ, *et al*. Immunohistological analysis of synovial tissue for differential diagnosis in early arthritis. *Rheumatology (Oxford)* 1999;**38**:1074–80.
167. **Vossenaar ER**, Smeets TJ, Kraan MC, *et al*. The presence of citrullinated proteins is not specific for rheumatoid synovial tissue. *Arthritis Rheum* 2004;**50**:3485–94.
168. **Mjaavatten MD**, Haugen AJ, Helgetveit K, *et al*. High anti-cyclic citrullinated peptide level is a stronger predictor than low level for persistent joint swelling

## Recommendations

- in patients presenting with arthritis of  $\leq 16$  weeks duration. *Arthritis Rheum* 2008;**58**:S770.
169. **Mjaavatten MD**, Nygaard H, Helgetveit K, *et al*. Clinical characteristics of patients presenting with oligoarthritis in a very early arthritis clinic in Norway: predictors of persistent arthritis at six month follow-up. *Arthritis Rheum* 2007;**56**:1638.
  170. **Baron M**, Schieir O, Hudson M, *et al*. The clinimetric properties of the World Health Organization Disability Assessment Schedule II in early inflammatory arthritis. *Arthritis Rheum* 2008;**59**:382–90.
  171. **Harwood RH**, Carr AJ, Thompson PW, *et al*. Handicap in inflammatory arthritis. *Br J Rheumatol* 1996;**35**:891–7.
  172. **Bykerk VP**, Mironyuk L, Chen H, *et al*. Validity of the RADAI in early rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**(Suppl II):344.
  173. **Baron M**, Steele R. Development of the McGill Range of Motion Index. *Clin Orthop Relat Res* 2007;**456**:42–50.
  174. **Bunn DK**, Shepstone L, Galpin LM, *et al*. The NOAR Damaged Joint Count (NOAR-DJC): a clinical measure for assessing articular damage in patients with early inflammatory polyarthritis including rheumatoid arthritis. *Rheumatology (Oxford)* 2004;**43**:1519–25.
  175. **Aletaha D**, Neogi T, Silman A, *et al*. 2010 Rheumatoid arthritis classification criteria. An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. *Ann Rheum Dis* 2010;**69**:1580–88.

## Chapter 3: Development and validation of a new index

Because no specific instrument can be recommended in early populations, we decided to develop and validate a new instrument in our cohort. We aimed to develop an index that needed no special calculation, had a normal distribution, accounted for sex differences, and resolved the issue of choosing between CRP level and ESR in multicenter observational studies and clinical practice.

Moreover, considering that early initiation of DMARDs adjusted to a tight control strategy is critical for the achievement of remission, we created an index in a mixed population of patients with early RA and patients with UA; both of these patient groups are regularly seen at early arthritis clinics.

The following article is presented:

Development and validation of a new disease activity index as a numerical sum of 4 variables in patients with early arthritis. (Arthritis Care Res (Hoboken) 2013 Apr;65(4):518-25)

**3.1. Development and validation of a new disease activity index as a numerical sum of 4 variables in patients with early arthritis. (Arthritis Care Res (Hoboken) 2013 Apr;65(4):518-25)**

# Development and Validation of a New Disease Activity Index as a Numerical Sum of Four Variables in Patients With Early Arthritis

ISABEL CASTREJÓN,<sup>1</sup> LORETO CARMONA,<sup>2</sup> ANA M. ORTIZ,<sup>1</sup> MIGUEL A. BELMONTE,<sup>3</sup>  
JUAN A. MARTÍNEZ-LÓPEZ,<sup>4</sup> AND ISIDORO GONZÁLEZ-ÁLVARO<sup>1</sup>

**Objective.** To describe the development and validation of a disease activity index in early arthritis that can be easily applied in daily practice and clinical research.

**Methods.** The Hospital Universitario La Princesa Index (HUPI) was developed after analysis of data from an early arthritis cohort (202 patients with 756 visits). It is the sum of 4 variables (graded 0–3): tender joint count, swollen joint count, patient global assessment, and acute-phase reactants (erythrocyte sedimentation rate [ESR] and/or C-reactive protein [CRP] level, depending on availability at the moment of evaluation). The score for each variable was based on its quartile distribution in the cohort. The HUPI was validated using the following properties: feasibility, internal consistency (Cronbach's alpha), convergent validity (Pearson's r coefficients with other activity measures), criterion validity (area under the receiver operating characteristic curve [AUC ROC] to detect minimal disease activity [MDA]), and sensitivity to change (AUC ROC) to detect change with the physician's and patient's assessment of disease activity.

**Results.** Internal consistency is reasonable ( $\alpha = 0.63$ ). The HUPI correlates well with activity measures such as the Disease Activity Score in 28 joints (DAS28;  $r = 0.89$ ) and the Simplified Disease Activity Index (SDAI;  $r = 0.70$ ), and correlates slightly worse with the functional index of the Health Assessment Questionnaire ( $r = 0.69$ ). It discriminates MDA correctly (AUC 0.95), and its sensitivity to change is slightly superior (AUC 0.902) to that of the DAS28-ESR (AUC 0.864), the DAS28-CRP (AUC 0.889), and the SDAI (AUC 0.791).

**Conclusion.** The HUPI has face validity, is easy to calculate, is sensitive, and is a valid composite index for the assessment of disease activity in patients with early arthritis, both in clinical research and in routine care.

## INTRODUCTION

Intensive management strategies based on enhanced assessment and tight control of disease activity can improve the outcome of patients with rheumatoid arthritis (RA) (1,2). Furthermore, the importance of early diagnosis is increasingly important, as recent-onset disease is more sensitive to treatment than later-stage disease. Conse-

quently, early initiation of treatment can slow or prevent disease progression.

Early arthritis clinics have been established in many countries and enable faster referral of patients with arthritis and earlier implementation of strategies to improve clinical and radiologic outcomes (3–6). The use of composite indices in these clinics has proven extremely useful in the followup and assessment of patients with early arthritis. Appropriately validated instruments are necessary in the daily practice. A task force, comprised of rheu-

Supported by a grant from Genoma España to the MEICA project (Molecular and Cellular Mechanisms in Chronic Inflammatory and Autoimmune Diseases) and by Fondo para la Investigación Sanitaria of the Instituto de Salud Carlos III (grant PI08/0754).

<sup>1</sup>Isabel Castrejón, MD (current address: NYU Hospital for Joint Diseases, New York), Ana M. Ortiz, MD, PhD, Isidoro González-Álvarez, MD, PhD: Hospital Universitario de La Princesa, Instituto de Investigación Sanitaria Hospital La Princesa, Madrid, Spain; <sup>2</sup>Loreto Carmona, MD, PhD: Universidad Camilo José Cela, Madrid, Spain; <sup>3</sup>Miguel A. Belmonte, MD, PhD: Hospital General de Castellón, Castellón, Spain; <sup>4</sup>Juan A. Martínez-López, MD: Fundación Jiménez Díaz, Madrid, Spain.

Dr. Ortiz has received consultant fees, speaking fees, and/or honoraria (less than \$10,000 each) from Abbott, Esteve, and MSD. Dr. Belmonte has received consultant fees, speaking fees, and/or honoraria (less than \$10,000 each) from Roche and MSD. Dr. González-Álvarez has provided expert testimony for Roche and UCB.

Address correspondence to Isidoro González-Álvarez, MD, PhD, Rheumatology Department, Hospital Universitario la Princesa, Calle Diego de León, 62, 28006 Madrid, Spain. E-mail: isidoro.ga@ser.es.

Submitted for publication November 28, 2011; accepted in revised form August 30, 2012.



## Significance & Innovations

- Because no single marker reflects all the characteristics of the disease, evaluation of disease activity in early rheumatoid arthritis (RA) and undifferentiated inflammatory arthritis is a challenge for rheumatologists.
- Although disease activity indices have significantly improved assessment of RA, they remain limited by their complexity and sex bias.
- We describe the validation of a new composite disease activity index based on the sum of 4 variables graded from 0 to 3: tender joint count, swollen joint count, patient global disease assessment, and erythrocyte sedimentation rate/C-reactive protein level. The index is feasible and sensitive to change, and could prove superior to previous indices in that it prevents bias arising from sex and missing data.
- This new index can be used in patients with early arthritis and will be validated in patients with late arthritis.

matologists and a patient, has recently developed a set of recommendations to use a treatment-to-target strategy in patients with RA. One of the recommendations was to apply this strategy to routine care (7). This specific recommendation found a high level of agreement between a large number of physicians representing 34 different countries (8). To be able to implement this treatment-to-target strategy, physicians would need to have a feasible and valid measure of disease activity to set a specific target.

The Disease Activity Score (DAS) was developed based on expert opinion as a measure of RA activity (9). Modifications of this score, i.e., the DAS in 28 joints (DAS28; simplified joint count) (10) and the DAS28 based on C-reactive protein (CRP) level instead of erythrocyte sedimentation rate (ESR) (11), were developed to provide a more accurate and feasible measure of disease activity. Although these versions have advantages over the DAS, they remain subject to limitations. Application of the DAS and its versions is complex, as it is necessary to use a DAS calculator. In addition, ESR thresholds differ for men and women, and women have higher tender joint counts (associated with a higher density of C-fiber nociceptors); therefore, the DAS28 may overestimate disease activity in women (12–15).

Considering CRP level as a more reliable acute-phase reactant (APR) than ESR (16), the DAS28-CRP could reduce sex bias. However, we previously showed that DAS28 results vary according to whether ESR or CRP level is used in the calculation (17). This finding could affect the results of multicenter observational studies in which using one marker or the other depends on local availability.

In the 2003 study by Smolen et al, the Simplified Disease Activity Index (SDAI), an unweighted and untransformed index (18), was developed; it is calculated as the sum of 5 variables, namely those included in the DAS28-CRP plus the physician's global disease assessment (GDA-Ph). These

same authors also developed the Clinical Disease Activity Index (CDAI), a composite index that does not incorporate an acute-phase response (19). The SDAI and CDAI results correlate closely with those of the DAS28, although they do not follow a normal distribution (considerable left shift compared with the DAS28). The absence of a normal distribution reduces the usefulness and interpretation of results in some statistical models when the SDAI score is the dependent variable. In addition, there is considerable disagreement in the classification into low and moderate disease activity between the DAS28 and the SDAI (20).

Consequently, we aimed to develop an index that needed no special calculator, had a normal distribution, accounted for sex differences, and resolved the issue of choosing between CRP level or ESR in multicenter observational studies and clinical practice. Therefore, considering that early initiation of disease-modifying antirheumatic drugs adjusted to a tight control strategy is critical for the achievement of remission, we created an index in a mixed population of patients with early RA and patients with undifferentiated arthritis (UA); both of these patient groups are regularly seen at early arthritis clinics.

## PATIENTS AND METHODS

We developed the Hospital Universitario de la Princesa Index (HUPI) using data from the Early Arthritis Registry of Hospital Universitario La Princesa (Madrid, Spain), which exhibits similar characteristics as compared with the Leiden Early Arthritis Clinic (Supplementary Table 1, available in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/acr.21854/abstract>), providing support for the generalizability of this population. Validation involved the use of followup data to analyze feasibility, validity, and sensitivity to change.

**Data source.** The Early Arthritis Registry of Hospital Universitario La Princesa includes all patients with  $\geq 2$  swollen joints for  $>4$  weeks and  $<1$  year and a diagnosis of RA (1987 revised criteria of the American College of Rheumatology [ACR]) (21) or undifferentiated inflammatory arthritis. The clinical protocol of the registry comprises 4 structured visits during a 2-year followup period and was reviewed and approved by the Ethics Committee for Clinical Research of Hospital Universitario La Princesa. Prior to inclusion in the register, all patients signed a written informed consent form.

Treatment was individualized and adapted as necessary based upon the decision of the treating rheumatologist. Routine clinical and laboratory data were collected as follows: rheumatoid factor, anti-cyclic citrullinated peptide antibody, ESR, CRP level, duration of symptoms, swollen joint count (SJC) and tender joint count (TJC) out of a total of 28 joints, the GDA-Ph and the patient's global disease assessment (GDA-P) on a 100-mm visual analog scale, and the Spanish version of the Health Assessment Questionnaire (HAQ) (22). The DAS28 with ESR or CRP level, the SDAI, and the CDAI are calculated automatically. Joint count assessments were performed by 2 experienced physicians (AMO and IG-A) to reduce interrater variability.

**Development of the HUPI.** Following the recommendations of Outcome Measures in Rheumatology for core set outcome measures (23,24) and the indices DAS (9), DAS28 (25), and SDAI (18), we based our index on the TJC, SJC, GDA-P, and the APR.

Each variable was divided into quartiles, each of which was assigned an ordinal value from 0 to 3 (see Table 1). Additionally, and based on evidence suggesting sex bias for the TJC and ESR (13,14), we defined different cutoff points for these variables stratified by sex (Table 1). As elderly patients exhibit higher GDA-P, we defined 2 strategies to assign the cutoffs for this variable. The first strategy was to divide the population into 2 groups depending on age ( $\leq 40$  years or  $> 40$  years); the resulting variable was called global disease assessment by patient by age (GDA-P1). The second strategy was not to divide the population, and the resulting variable was called global disease assessment by patient irrespective of age (GDA-P2).

The CRP level was scored using 2 strategies: one according to quartile distribution (CRP1) and the other according to theoretical thresholds based on local reference ranges (CRP2).

In addition to the development of the HUPI versions that included only ESR, CRP1 level, or CRP2 level, we described 4 different possibilities to input the APR. Versions with APR1 and APR2 only were calculated if the CRP level and ESR were both available. In this case, the APR1 approach was to use the average of the scores of the ESR and CRP1 level, and the APR2 approach was to use the average of the scores of the ESR and CRP2 level. Versions including APR3 and APR4 were calculated with the scores of ESR or CRP level (CRP1 level with APR3, CRP2 level with APR4) when only one of them was available or with the average of their scores when both were available.

We developed 10 alternative HUPI versions, each of which was the sum of the scores assigned to each of these 4 stratified variables (Supplementary Table 2, available in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/acr.21854/abstract>).

Each variable was discrete, and the total ranged from 0 to 12. Most variables had increments of 1 point between levels, although some had increments of 0.5 point (i.e., those including APR1–4; see Supplementary Table 2, available in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/acr.21854/abstract>).

**Validation and selection of the best index.** To determine which of these versions was the most reliable and useful, we analyzed different aspects of the validation process, namely feasibility, reliability, construct validity, and responsiveness.

Feasibility includes domains such as completion time, difficulty, clarity, and acceptance by both patients and clinicians. We quantified feasibility by creating an ad hoc measure ranging from 0 (unfeasible) to 3 (completely feasible) to evaluate 3 domains: completion time (according to the number of variables included); clarity of the calculation (depending on the simplicity of the variables); and acceptance (low probability of missing data). Each author independently rated each index in the 3 domains. The final rating of the versions was the mean of the 3 values.

Reliability embraces the concepts of internal consistency and reproducibility. The internal consistency or “good construction” of each HUPI version was tested using Cronbach’s alpha (where  $\alpha = < 0.70$  indicates that individual items provide an inadequate contribution to the overall scale, and values of  $\alpha = > 0.90$  suggest redundancy).

Construct validity refers to the proximity of our measure to similar measures (convergent validity) and distance from dissimilar measures (divergent validity). When comparing the HUPI with similar construct measures (disease activity measures), a high correlation (Pearson’s  $r$ ) would be expected; when comparing it with less closely related constructs, such as function, a lower correlation would be expected. We tested the HUPI against the DAS28 and the SDAI, and then against the HAQ.

**Table 1. Scoring of the individual components of the HUPI, based on the distribution of the variables by quartiles\***

Variable	0	1	2	3
Swollen joint count	0	1–2	3–4	$> 4$
Tender joint count				
Male	0	1	2–3	$> 3$
Female	0	1–2	3–6	$> 6$
GDA-P1†				
Ages $\leq 40$ years	0–10	11–20	21–40	$> 40$
Ages $> 40$ years	0–20	21–40	41–50	$> 50$
GDA-P2	0–15	16–30	31–50	$> 50$
CRP1 level, mg/dl†	$\leq 0.30$	0.31–0.50	0.51–1	$> 1$
CRP2 level, mg/dl	$\leq 0.10$	0.11–0.80	0.81–1.50	$> 1.50$
ESR				
Male	0–10	11–15	16–20	$> 20$
Female	0–15	16–20	21–30	$> 30$

\* HUPI = Hospital Universitario de la Princesa Index; GDA-P1 = global disease assessment by age; GDA-P2 = global disease assessment by patient irrespective of age; CRP1 = C-reactive protein level by quartile distribution; CRP2 = CRP level by theoretical thresholds based on local reference ranges; ESR = erythrocyte sedimentation rate.

† Excluded because category did not improve the validity of the index.



Criterion validity was evaluated using receiver operating characteristic (ROC) curves with minimal disease activity (MDA) (26) as the external criterion. MDA was developed in 2005 by Wells et al (26) as a satisfactory state of disease activity to compare different treatment strategies, bearing in mind that true remission is difficult to achieve in patients with RA. Two equivalent definitions were formulated, one based on the DAS28 (European League Against Rheumatism [EULAR] response criteria) and the other based on meeting cutoffs in 5 of the 7 World Health Organization/International League of Associations for Rheumatology core set outcome measures, which is the set used in our analysis. The statistic applied was the area under the curve (AUC) (27), and the ROC curve of the HUPI was compared with that of the DAS28-ESR, the DAS28-CRP, and the SDAI using the *roccomp* command of Stata.

Responsiveness, also called sensitivity to change, is defined as “the ability of an instrument to accurately detect change when it has occurred” (28), implying that the intervention administered to the study patients involved an effect with a known direction. In our cohort, the intervention was the treatment initiated by the physician, which in most instances was methotrexate. We analyzed the AUC of the change in the HUPI for identifying patients who improved after 6 months of treatment (29). Responsiveness was tested against 3 definitions of improvement as follows: 1) a change in the GDA-Ph  $>10$  between baseline and 6 months of followup; 2) the same definition but for GDA-P; and 3) change in the DAS28 compared with the change in

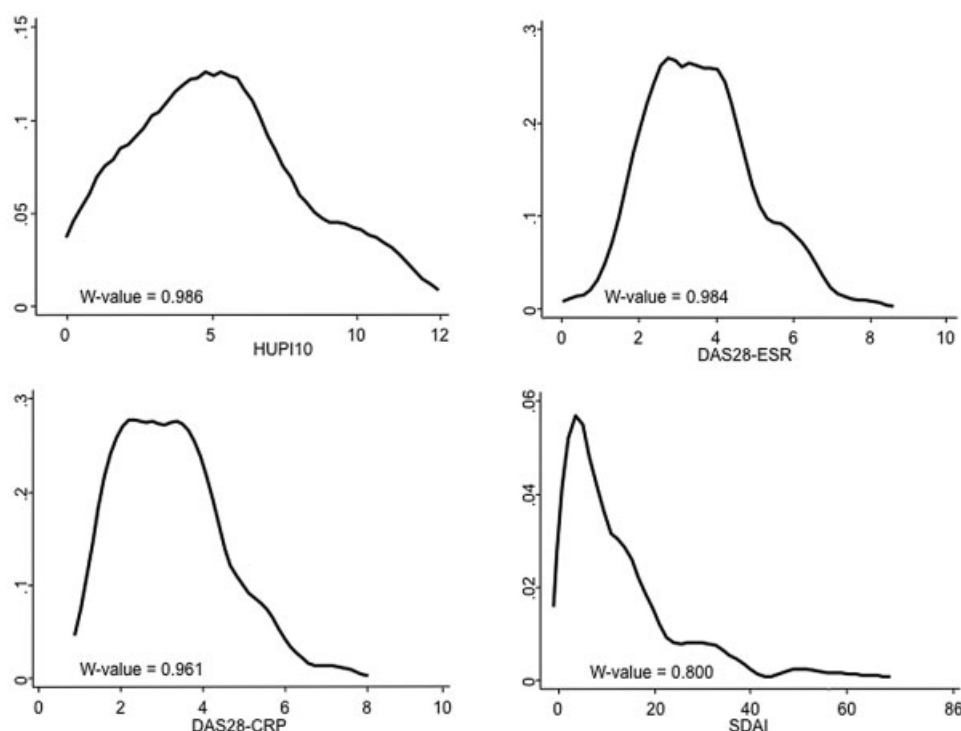
the HUPI. Responsiveness by the first 2 definitions was tested with the AUC of the ROC curves using the *roccomp* command of Stata to determine statistically significant differences between the different indices. Responsiveness by the third definition was tested with the beta coefficient from the linear regression analysis.

Statistical significance was set at a *P* value of less than 0.05; if Bonferroni correction because of multiple comparisons was needed, then the *P* value was set at less than 0.0125.

## RESULTS

We analyzed 756 visits corresponding to 202 patients (2–4 visits/patient, mean 3.6 visits/patient), of whom 77% were women. Mean  $\pm$  SD age at onset was  $53 \pm 16$  years. At the end of followup, 70% fulfilled ACR 1987 revised criteria (21) for RA, and 30% were classed as having UA. A more detailed description of this population has been published previously (17).

None of the 10 versions of the HUPI exhibited a perfect Gaussian distribution, although the values obtained in the Shapiro-Wilk test were similar to those obtained for the 2 versions of the DAS28 and were slightly higher than those obtained for the SDAI (Figure 1 and Table 2). In addition, all 10 versions exhibited comparable validity (Table 2), although HUPI version 10 fared better in most of the validity aspects and was thus selected. All further references to the HUPI concern this version of the index.



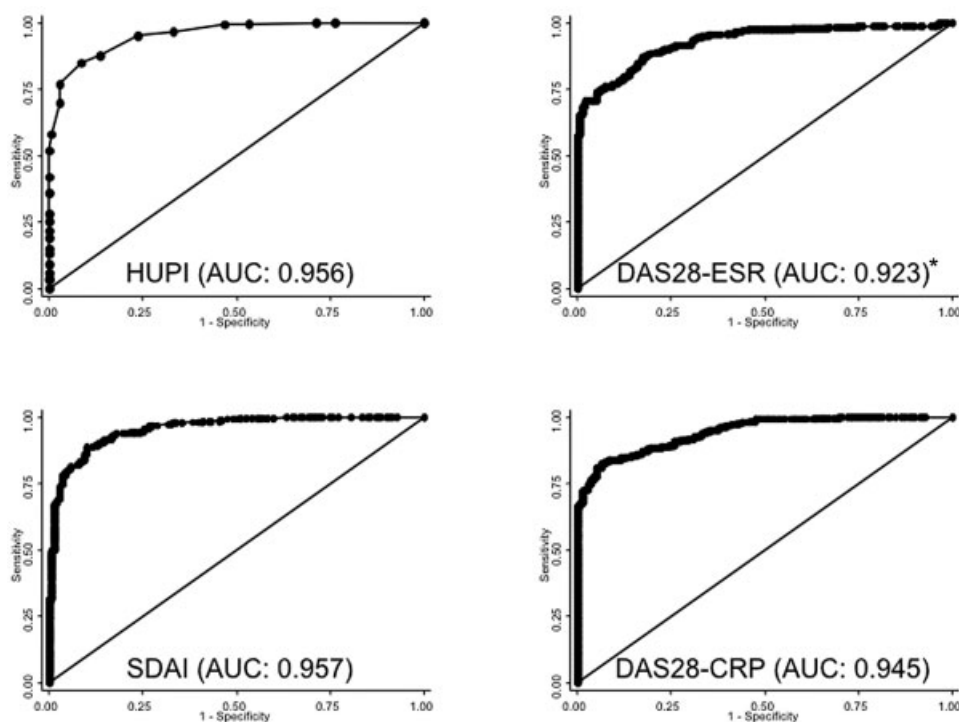
**Figure 1.** Distribution of the values for the Hospital Universitario de la Princesa Index, version 10 (HUPI10), the Disease Activity Score in 28 joints (DAS28) with erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) level, and the Simplified Disease Activity Index (SDAI) in the entire cohort ( $n = 202$ ). The W value refers to the score obtained by each index in the Shapiro-Wilk normality test; higher values represent a better fit to the normal distribution.

Table 2. Validation of the 10 HUPI versions\*

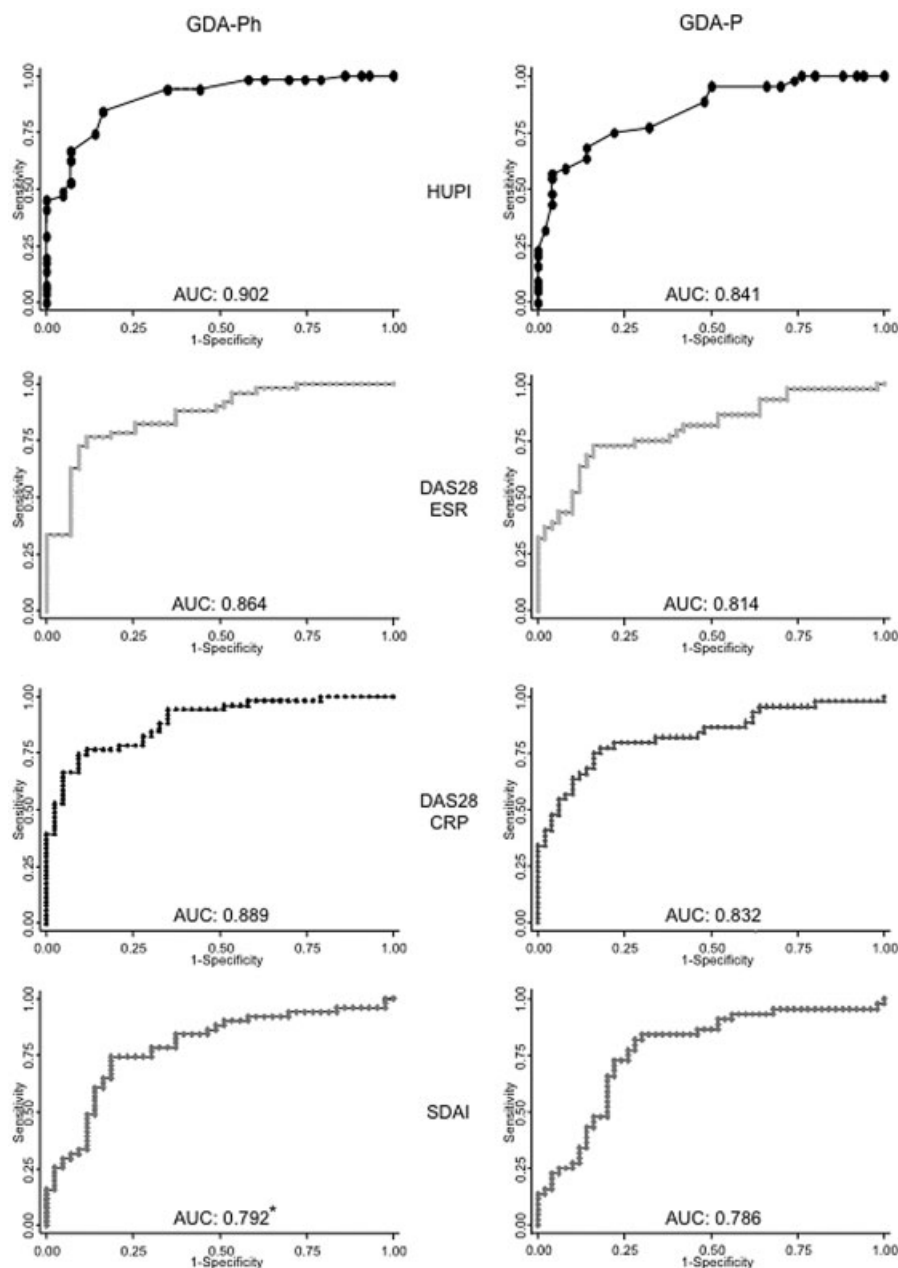
Feasibility, T+E+A		Construct validity			Criterion validity, MDA	Normality, W score†	Responsiveness	
		DAS	SDAI	HAQ			GDA	DAS28
HUPI version								
1	2	0.900‡	0.712‡	0.673	0.944	0.990‡	0.965‡	0.931‡
2	2	0.845	0.684	0.646	0.934	0.983	0.944	0.920
3	1.2	0.884	0.708‡	0.661	0.950	0.984	0.947‡	0.949‡
4	2.4‡	0.899‡	0.702	0.676	0.946	0.992‡	0.947‡	0.912
5	2.4‡	0.846	0.676	0.650	0.938	0.986	0.902	0.907
6	2.4‡	0.849	0.694	0.670	0.948	0.982	0.911	0.908
7	1.4	0.884	0.699	0.665	0.952‡	0.986‡	0.916	0.933‡
8	1.4	0.886‡	0.708‡	0.676‡	0.956‡	0.985	0.921	0.895
9	1.3	0.884	0.699	0.678‡	0.952‡	0.986	0.946	0.899
10	1.9	0.886‡	0.708‡	0.687‡	0.956‡	0.986‡	0.948‡	0.898
* HUPI = Hospital Universitario de la Princesa Index; T = completion time; E = ease of calculation with regard to the different items included; A = acceptance by the patient and the physician; DAS = Disease Activity Score; SDAI = Simplified Disease Activity Index; HAQ = Health Assessment Questionnaire; MDA = minimal disease activity; GDA = global disease assessment; DAS28 = DAS in 28 joints.								
† By Shapiro-Wilk normality test.								
‡ One of 3 higher values for each validation aspect.								

**Construct validity.** The mean  $\pm$  SD HUPI score was  $6.51 \pm 3.18$  at baseline and  $4.35 \pm 2.59$  at 6 months. The HUPI was calculated at 722 of 756 visits because of missing data, whereas the DAS28-ESR was only calculated at 684 visits and the DAS28-CRP and SDAI at 664 visits. As expected, the HUPI score correlated with the variables that

measure disease activity ( $\rho = 0.89$  and  $n = 684$  for DAS28-ESR;  $\rho = 0.91$  and  $n = 664$  for DAS28-CRP;  $\rho = 0.71$  and  $n = 664$  for SDAI; and  $\rho = 0.82$  and  $n = 664$  for CDAI) and the correlation between the HUPI and the HAQ was also high ( $\rho = 0.69$ ). The same was true for the other disease activity indices (DAIs; data not shown).



**Figure 2.** Receiver operating characteristic curves to compare the accuracy of the Hospital Universitario de la Princesa Index (HUPI), version 10, compared with that of the Disease Activity Score in 28 joints (DAS28) with C-reactive protein (CRP) level and erythrocyte sedimentation rate (ESR) and the Simplified Disease Activity Index (SDAI), using minimal disease activity as a reference. The area under the curve (AUC) is shown at the bottom of each curve. The roccomp command of Stata was used to determine statistically significant differences between the different indices. \* = statistical significance with respect to HUPI was set at  $P < 0.0125$  (with Bonferroni correction for multiple comparisons).



**Figure 3.** Comparison of the responsiveness of the Hospital Universitario de la Princesa Index (HUPI), version 10, with the Disease Activity Score in 28 joints (DAS28) with C-reactive protein (CRP) level and erythrocyte sedimentation rate (ESR), and the Simplified Disease Activity Index (SDAI) using 2 different external criteria: improvement based on physician global disease assessment (GDA-Ph) at baseline and the 6-month visit (left column) and improvement based on patient global disease assessment (GDA-P) at baseline and the 6-month visit (right column). The area under the curve (AUC) is shown at the bottom of each curve. The roccomp command of Stata was used to determine statistically significant differences between the different indices. \* = statistical significance with respect to the HUPI was set at  $P < 0.0125$  (with Bonferroni adjustment for multiple comparisons).

**Reliability.** Although internal consistency was modest for the HUPI, with a Cronbach's alpha of 0.63, it was still better than for the DAS28-ESR ( $\alpha = 0.52$ ), the DAS28-CRP ( $\alpha = 0.47$ ), the SDAI ( $\alpha = 0.48$ ), and the CDAI ( $\alpha = 0.46$ ). We did not specifically test reproducibility, since the reproducibility of joint counts and global disease assessment is not very good (30).

**Criterion validity.** To compare how the 5 DAIs discriminate MDA, we used only the 664 visits in which all 5 indices were estimated. The AUC for the HUPI was very high (0.956), slightly larger than that of the DAS28-ESR (0.930;  $P = 0.001$ ) or the DAS28-CRP (0.945;  $P = 0.077$ ), and similar to that of the SDAI (0.957;  $P = 0.971$ ) (Figure 2) and the CDAI (0.964;  $P = 0.343$ ) (data not shown).

**Sensitivity to change.** Using the physician assessment and the patient assessment as external criteria of change, we compared the responsiveness of the HUPI after 6 months of treatment with that of the DAS28-ESR, the DAS28-CRP, the SDAI, and the CDAI based on data from the 94 patients for whom the information of all DAIs was available at baseline and at the second visit. With GDA-Ph, the AUC for the HUPI (0.902) was slightly larger than that of the DAS28-ESR (0.864;  $P = 0.229$ ), DAS28-CRP (0.889;  $P = 0.625$ ), SDAI (0.792;  $P = 0.01$ ), and CDAI (0.791;  $P = 0.002$ ). With GDA-P, the AUC for HUPI (0.841) was similar to that of the DAS28-ESR (0.814;  $P = 0.218$ ), DAS28-CRP (0.833;  $P = 0.739$ ), and SDAI (0.786;  $P = 0.208$ ) and was slightly lower than the one for CDAI (0.987;  $P < 0.001$ ).

Responsiveness was also tested using linear regression analysis. The difference in the DAS28-ESR from baseline to 6 months was the dependent variable. The  $\beta$  coefficient for the difference in HUPI was 0.85.

## DISCUSSION

Measurement of disease activity is considered a standard approach in clinical practice that facilitates management and followup of patients. Reliable evaluation has been possible using the RA core set variables (31,32) and composite indices described in the literature, the most well-known being the DAS28 (10), the SDAI (18), and the CDAI (33). These indices are difficult to apply because of the complexity of calculation, sex and age bias (DAS28) (12–14,34,35), and a nonweighted design shifting the distribution to the left (SDAI and CDAI) (Figure 1; data not shown for CDAI). In order to overcome these drawbacks, we developed and validated a new index, the HUPI, which includes the same variables used in the DAS28 and the SDAI. The HUPI is simple to calculate and seems at least as accurate and sensitive to change as previously validated indices.

The HUPI was developed by analyzing the balance between simplicity, reliability, accuracy, and sensitivity. In the resulting versions, the variables were weighted according to their quartile distribution in the study population. The 10 versions of the index varied with the weighting applied to the different variables (e.g., sex, age, APR cutoffs, and application of CRP level or ESR). The best index, the HUPI10, used a common score for SJC and GDA-P and a sex-adjusted score for TJC and ESR. Following a common approach in large multipractice registries, we applied ESR, CRP level, or both, depending on availability.

Since the HUPI can be completed within a few minutes, it is more suited to daily clinical practice than the DAS28. It is not as easy to calculate as the SDAI, since the calculation requires a table with the different cutoffs, although given the left shift in the SDAI distribution, the sensitivity to change of the HUPI is significantly better than that of the SDAI and slightly higher than both DAS28-ESR and DAS-CRP.

The key advantage of the HUPI is the possibility of using ESR, CRP level, or both as the APR. Missing data on ESR or CRP level is a frequent problem in observational studies. DAS28 values estimated with ESR or CRP level are not

equivalent (17,36), as occurs with the SDAI and CDAI (19). Furthermore, CRP level is more effective and informative than ESR in some patients and vice versa. Therefore, the possibility of using both APRs, and the fact that HUPI values cover the complete range of the index (whereas values of the other DAIs only span part of their ranges [Figure 1]), may account for its sensitivity to change. Consequently, application of the HUPI in clinical trials might help to reduce the number of patients needed to test differences between comparators.

In contrast to our work, most studies on validation have been performed in patients with established RA. Our cohort includes patients with early RA and UA. This approach is consistent with the current trend of early management and diagnosis of RA that led to the development of the new ACR/EULAR 2010 classification criteria (37). To our knowledge, only one other publication has tried to address the validation of DAS in UA (38). Our group is performing an additional validation in longstanding disease.

Our study is limited by the fact that validation of an instrument is an ongoing process. In addition, we tested validity in a single population using a single data set; therefore, our results cannot be extrapolated to other populations. Although, we reproduced very similar results when testing it in another early arthritis data set (data not shown). The thresholds we describe for the variables included in the index may require fine-tuning once the HUPI is validated in different cohorts, especially in the case of long-term RA. It would also be interesting to evaluate whether sensitivity to change is influenced by genetic or sociocultural backgrounds, or by differences in measurement of ESR and CRP level, which also affect currently used indices. Future objectives will include the development of thresholds for HUPI to distinguish remission based on recently published criteria (39) and thresholds for low, moderate, and high activity.

In summary, we provide evidence that the HUPI is feasible and sensitive to change in disease activity. In addition, it is accurate and makes it possible to avoid bias arising from sex and missing data.

## AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. González-Álvaro had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study conception and design.** Castrejón, Carmona, Belmonte, González-Álvaro,

**Acquisition of data.** Castrejón, Ortiz, González-Álvaro,

**Analysis and interpretation of data.** Castrejón, Carmona, Martínez-López, González-Álvaro.

## REFERENCES

1. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364:263–9.
2. Gonzalez-Alvaro I, Descalzo MA, Carmona L. Trends towards an improved disease state in rheumatoid arthritis over time: influence of new therapies and changes in management ap-



- proach. Analysis of the EMECAR cohort. *Arthritis Res Ther* 2008;10:R138.
3. Descalzo MA, Carbonell J, Gonzalez-Alvaro I, Sanmarti R, Balsa A, Hernandez-Barrera V, et al, on behalf of the SERAP and PROAR Study Groups. Effectiveness of a clinical practice intervention in early rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2012;64:321–30.
  4. Sokka T, Pincus T. A historical perspective concerning population-based and clinical studies of early arthritis and early rheumatoid arthritis. *Clin Exp Rheumatol* 2003;21:S5–14.
  5. Combe B, Benessiano J, Berenbaum F, Cantagrel A, Daures JP, Dougados M, et al. The ESPOIR cohort: a ten-year follow-up of early arthritis in France. Methodology and baseline characteristics of the 813 included patients. *Joint Bone Spine* 2007;74:440–5.
  6. De Rooy DP, van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH. Predicting arthritis outcomes: what can be learned from the Leiden Early Arthritis Clinic? *Rheumatology (Oxford)* 2011;50:93–100.
  7. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
  8. Haraoui B, Smolen JS, Aletaha D, Breedveld FC, Burmester G, Codreanu C, et al. Treating rheumatoid arthritis to target: multinational recommendations assessment questionnaire. *Ann Rheum Dis* 2011;70:1999–2002.
  9. Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49:916–20.
  10. Van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41:1845–50.
  11. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. *Ann Rheum Dis* 2009;68:954–60.
  12. Ahlmen M, Svensson B, Albertsson K, Forslind K, Hafstrom I. Influence of gender on assessments of disease activity and function in early rheumatoid arthritis in relation to radiographic joint damage. *Ann Rheum Dis* 2010;69:230–3.
  13. Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther* 2009;11:R7.
  14. Castrejon Fernandez I, Martinez-Lopez JA, Ortiz Garcia AM, Carmona Ortells L, Garcia-Vicuna R, Gonzalez-Alvaro I. Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid. *Reumatol Clin* 2010;6:134–40.
  15. Unruh AM. Gender variations in clinical pain experience. *Pain* 1996;65:123–67.
  16. Kushner I. C-reactive protein in rheumatology. *Arthritis Rheum* 1991;34:1065–8.
  17. Castrejon I, Ortiz AM, Garcia-Vicuna R, Lopez-Bote JP, Humbria A, Carmona L, et al. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? *Clin Exp Rheumatol* 2008;26:769–75.
  18. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003;42:244–57.
  19. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23:S100–8.
  20. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T, et al. Disease activity measurement of rheumatoid arthritis: comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum* 2005;53:56–60.
  21. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
  22. Esteve-Vives J, Batlle-Gualda E, Reig A. Spanish version of the Health Assessment Questionnaire: reliability, validity and transcultural equivalency. Grupo para la Adaptacion del HAQ a la Poblacion Espanola. *J Rheumatol* 1993;20:2116–22.
  23. Tugwell P, Boers M. Developing consensus on preliminary core efficacy endpoints for rheumatoid arthritis clinical trials: OMERACT Committee. *J Rheumatol* 1993;20:555–6.
  24. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction. *J Rheumatol* 1993;20:528–30.
  25. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
  26. Wells GA, Boers M, Shea B, Brooks PM, Simon LS, Strand CV, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;32:2016–24.
  27. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986;39:897–906.
  28. De Bruin AF, Diederiks JP, de Witte LP, Stevens FC, Philipsen H. Assessing the responsiveness of a functional status measure: the Sickness Impact Profile versus the SIP68. *J Clin Epidemiol* 1997;50:529–40.
  29. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
  30. Scott DL, Housien DA. Joint assessment in rheumatoid arthritis. *Br J Rheumatol* 1996;35 Suppl:14–8.
  31. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;36:729–40.
  32. Scott DL, Panayi GS, van Riel PL, Smolen J, van de Putte LB. Disease activity in rheumatoid arthritis: preliminary report of the Consensus Study Group of the European Workshop for Rheumatology Research. *Clin Exp Rheumatol* 1992;10:521–5.
  33. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7:R796–806.
  34. Leeb BF, Haindl PM, Maktari A, Nothnagl T, Rintelen B. Disease activity score-28 values differ considerably depending on patient's pain perception and sex. *J Rheumatol* 2007;34:2382–7.
  35. Radovits BJ, Fransen J, van Riel PL, Laan RF. Influence of age and gender on the 28-joint Disease Activity Score (DAS28) in rheumatoid arthritis. *Ann Rheum Dis* 2008;67:1127–31.
  36. Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. *Ann Rheum Dis* 2007;66:407–9.
  37. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580–8.
  38. Fransen J, Visser K, van Dongen H, Huizinga T, van Riel PL, van der Heijde DM. Validity of the disease activity score in undifferentiated arthritis. *Arthritis Care Res (Hoboken)* 2010;62:1392–8.
  39. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;70:404–13.

## **PART III – GENERAL DISCUSSION**

### **Summary of results**

Our first study shows that the DAS28-ESR and DAS28-CRP values may not be interchangeable, as DAS28-ESR tends to yield higher values in women and long-standing RA patients (**Chapter 1.1.**). Although the correlation coefficient of the two indexes was high (0.91), the value of DAS28-ESR was higher than DAS28-CRP in approximately 90% of the visits. In addition, in a fourth of the visits analysed the differences between both indexes were greater than 1.2 points. This difference represent the minimal improvements considered to be relevant by EULAR criteria depending on baseline DAS28 measurement (van Gestel, Haagsma et al. 1998). When using the cut-off points proposed by Prevoo *et al.* (Prevoo, van 't Hof et al. 1995) different percentages of patients on remission were found: 41% when applying DAS28-CRP, and 26% when applying the DAS28-ESR. Using interchangeably DAS28-ESR or DAS28-CRP with the proposed cut-off values, we might estimate disease activity wrong and overestimate the proportion of patients in remission or with low disease activity. Indeed, the estimated cut-off values proposed by Inoue *et al.* for DAS28-CRP (Inoue, Yamanaka et al. 2007) support our finding that the DAS28-CRP values are, on average, 0.5 points lower than the DAS28-ESR values. Both indices could be better used in different situations. For example, DAS-CRP would be more useful in clinical trials, where we desire an index that showed the effect of therapeutic agents rapidly, and CRP responds faster than ESR. In contrast, DAS28-ESR may be more useful in routine care, where to know the patient's disease activity along the last months is preferable. Moreover, we demonstrated that DAS28-ESR is less sensitive to changes in long-standing disease and, in women the frequent higher values of ESR compared to men, may lead to a lower rate of remission as per the DAS28-ESR definition. These first results indicate that when the DAS28 is calculated with the CRP it may be more accurate to determine RA activity, especially in long-standing female patients.

After establishing that DAS28-CRP produces lower values than DAS28-ESR and knowing that other authors, as Inoue et al., have proposed preliminary cut-off values for DAS28-CRP (Inoue, Yamanaka et al. 2007) we then decided to described new cut-off values in our population (**Chapter 1.2.**). Our data suggest that DAS28 might behave similarly when applied to RA or UA patients. For this reason and also because it would be useful to have a validated composite

index for UA populations we aim to develop the specific cut-off point for DAS28-CRP in our mixed population. Moreover, comparisons between different populations should provide additional information about the reliability and reproducibility of these indexes. To estimate cut-off points for the DAS28-CRP in our cohort we established as a gold standard of disease activity the evaluation by six rheumatologists based on TJC, SJC, PGA, HAQ, ESR, and CRP. The best agreement between physicians was observed for the state of remission ( $\kappa = 0.65$ ), followed by high activity ( $\kappa = 0.52$ ). In contrast, there was only moderate concordance in the intermediate degrees of disease activity ( $\kappa = 0.30\text{--}0.41$ ). These values reflect the variability between physicians in their perception of disease activity. In a multivariate logistic regression model, TJC and HAQ contributed to the disagreement. Interestingly, evaluators' characteristics, such as gender, age, or years of experience, did not contribute to the disagreement in the perception of disease activity. The cut-off points for DAS28-CRP were obtained using ROC curve analysis. The best cut-off values of the DAS28-CRP to stratify the patients in our population according to the state of their disease were 2.3 for remission-low (Se 87%; Sp 96%), 3.8 for low-moderate (Se 78%; Sp 88%), and 4.9 for moderate-high disease activity (Se 84%; Sp 83%). These differed from the cut-off points obtained for DAS28-ESR with 2.7, 4.3, and 5.5, respectively. The AUC were always greater for DAS28-CRP than for DAS28-ESR at each level of disease activity, although statistically these differences were only significant for the low/moderate activity group. A second approach to estimate a set of cut-off points for DAS28-CRP was to use the conventional DAS28-ESR cut-off points as gold standard. Although this strategy did not allow us to estimate which index would be more accurate, we obtained the following cut-off values: 2.35 for remission-low activity (Se 90%; Sp 92%), 2.95 for low-moderate activity (Se 91.5%; Sp 90.5%), and 4.35 for moderate-high activity (Se 93%; Sp 96%). A more practical approach for daily clinical practice may be to use the MDA rather than the four levels of disease activity. The best MDA threshold estimated for DAS28-ESR was 2.8 (Se 86%; Sp 83%), and for DAS28-CRP 2.6 (Se 85%; Sp 89%). The AUC was again significantly higher for DAS28-CRP than for DAS28-ESR, suggesting that, to assess MDA, the DAS28-CRP cut-off point may be more accurate. Our best MDA thresholds for both DAS28-ESR and DAS28-CRP were very close to their respective cut-off values for remission, suggesting that MDA and clinical remission are very similar concepts.

When we explored the influence of gender on treatment response in our cohort we saw that, before initiating treatment, women showed higher levels of disease activity and disability compared to men, as reflected by higher values of DAS28 and HAQ (**Chapter 1.3.**). The ESR showed a tendency to be higher in women. However, there was no gender difference in the percentage of patients with poor prognostic factors present, such as RF, anti-CCP, or time to first DMARD. As a result of a higher level of disease activity, women received more aggressive treatment than men, reflected in higher percentage of combination therapy and in higher mean doses of methotrexate. Furthermore, the percentage of women to whom leflunomide was prescribed was twice as many as in men; TNF-blocking agents were used more often in women than in men as well. Although women received initially more aggressive treatment than men did, their DAS28-ESR remained higher than in men. The gender difference on DAS28-CRP was less pronounced once treatment was established. One explanation is that despite treatment, women had a lower treatment response and maintained a true higher disease activity. To test

this possibility we analysed global assessment by patient and by physician in both genders. Both global assessments were higher in women patients before treatment, but showed no differences by gender in subsequent visits once treatment was initiated. Concerning other items included in the DAS28-ESR, women showed higher TJC than men during follow-up. By gender, not only a difference in treatment response exists, but also a difference in TJC and ESR, all of which may influence the DAS28-ESR values. The impact of this gender difference in the evaluation of disease activity by DAS28 is of great consequences for clinical trials (Aletaha, Landewe et al. 2008). Not only women reach remission less frequently than men; they start with higher DAS28 and present worst response rates according to the EULAR criteria (Prevoo, van Gestel et al. 1996). As demonstrated in this study, DAS28-CRP has less potential bias in the assessment of disease activity, although certain gender differences persist due to the high weight of the TJC. Because the evaluation of treatment response rates based on currently available indices has a potential gender bias, it would be important to develop new indices that avoid this bias and provide a more objective assessment when making treatment decisions and evaluating the results in clinical trials.

Another concern about the use of these composite indices is that, although the use of such indices has become an important aspect of the care for patients with RA, we are unsure if they are equally useful for patients with UPIA. As we previously described, UPIA is a form of arthritis that does not fulfil classification criteria for a more definitive diagnosis. Due to the lack of a more precise clinical picture and outcome, it is important to have comprehensive tools that help the clinician anticipate outcomes. Studies focused on UPIA have used a myriad of indices to evaluate outcome, and we wanted to analyse whether they are valid to use in this population. After a systematic literature search, we found five studies evaluating the validity of different clinical measures in patients with UPIA (**Chapter 2**). Validation aspects of four questionnaires (WHO Disability Assessment Schedule, London Handicap Scale, Disease Repercussion Profile and the HAQ) and three physical measures (RA Disease Activity Index, McGill Range of Motion Index and NOAR Damage Joint Count) were partially assessed in these studies. We found no study on the validation of the most common activity measures, such as the DAS8 or the SDAI, in patients with UPIA. Although we found no direct evidence on what was the most useful index to follow-up patients with UPIA, experts decided, in light of very limited evidence, and mainly based on their experience, that disease activity should be monitored with whatever instrument, as none could be recommended.

The final part of the thesis was to describe the development and validation of a disease activity index to be applied in patients with early arthritis in daily practice and clinical research (**Chapter 3**). We initially described 10 versions of a new composite index called HUPI by analysing the balance between simplicity, reliability, accuracy, and sensitivity. In the resulting versions, the variables were weighted according to their quartile distribution in the study population. All versions exhibited a Gaussian distribution and comparable validity. Because HUPI version 10 fared better in most of the validity aspects, this was the version selected. HUPI-10 is calculated as the simple addition of the TJC, the number of SJC, the PGA and ESR, or CRP or  $(ESR+CRP)/2$ . We explored different aspects of the validity of the HUPI final version. Concerning construct



validity this new composite index showed a high correlation with the variables that measure disease activity, such as DAS28-ESR ( $\rho=0.89$ ); DAS28-CRP ( $\rho=0.91$ ); SDAI ( $\rho=0.71$ ); and CDAI ( $\rho=0.82$ ). The correlation with a measure for physical function as HAQ was slightly lower ( $\rho=0.69$ ). Although its internal consistency was modest, with a Cronbach's alpha of 0.63, it was higher than for the other composite indices, which had a Cronbach's alpha ranging from 0.46 to 0.52. To explore criterion validity we compared how the five disease activity indices discriminated MDA and all of them exhibited similar AUCs. Responsiveness was tested using the physician assessment and the patient assessment as external criteria of change, and using linear regression analysis. With both methods we reached similar results, the sensitivity to change of the HUPI was significantly better than that of the SDAI and slightly higher than both DAS28-ESR and DAS-CRP. Since the HUPI can be completed within a few minutes, it is more suited to daily clinical practice than the DAS28 or SDAI. The key advantage of the HUPI is the possibility of using ESR, CRP level, or both avoiding the problem with missing data when using laboratory tests.

## Strengths and limitations of the studies

It is important to consider that most of our analyses are based on a population where only 60% of patients fulfil the ACR criteria for RA classification, being the remaining patients undifferentiated arthritis (Arnett, Edworthy et al. 1988). Although this might introduce bias, we saw that both DAS28-ESR and DAS28-CRP behave similarly in both these subgroups of patient. Moreover, this mixed population provides the opportunity of evaluating these composite indexes in undifferentiated arthritis.

Other limitation is that factors potentially associated with changes in ESR or CRP, such as race, smoking, increased blood pressure, diabetes, high body mass index or abdominal adiposity, which could be relevant in our analysis, are not routinely collected in this cohort (Khera, McGuire et al. 2005). We did not adjust for all these variables since this information was not available, so we cannot exclude that they may influence the results.

Another limitation in our study is the possibly inadequacy of the gold standard used. To generate the cut-off values for DAS28-CRP we used the evaluation of disease activity by six rheumatologists as the gold standard. These may not represent a strong enough gold standard for two reasons, 1) the small number of rheumatologists involved in the evaluation, and 2) the fact that all the evaluators belonged to the same rheumatology department. Even having such a small sample from the same department there was some disagreement between physicians, especially in the moderate range of disease activity. Hence, individual variation in the perception of intermediate levels of disease activity is likely to exist.

In the second part of the thesis, and although we performed an exhaustive literature review to identify composite measures validated in undifferentiated populations, we found no study on the validation of the most common activity measures, so the experts decided in light of very limited evidence and mainly based on their experience.

Another limitation in our systematic review is that the study of validity of any clinical measure in undifferentiated arthritis is very challenging, in particular since “disease activity” implies that a defined disease should be diagnosed. Further, many of the included studies do not evaluate all aspects of the validation for an instrument.

Although we did our best effort to develop and validate an index to be used in early and undifferentiated arthritis, our last study is limited by the fact that validation of an instrument is an on-going process. In addition, we tested validity in a single population using a single data set; therefore, our results cannot be totally extrapolated to other populations until it is reproduced; the HUPI needs to be tested in different cohorts, including long-term RA as well.

An important strength in our work is that it is in agreement with previous published studies and supports previous results indicating that the cut-off points for DAS28-CRP and DAS28-ESR are not exchangeable. In addition, our data suggest that DAS28-CRP is more accurate. But, further studies would be necessary to confirm whether these cut-off points for DAS28-CRP can be applied in other populations.

Another strength is that most studies on validation have been performed in patients with established RA. Our cohort includes patients with early RA and undifferentiated arthritis. This approach is consistent with the current trend of early management and diagnosis of RA that led to the development of the new ACR/EULAR 2010 classification criteria (Aletaha, Neogi et al. 2010). To our knowledge, only one other publication has tried to address the validation of DAS in UA (Fransen, Visser et al. 2010)

## **Toward improved measures in early and undifferentiated arthritis**

Over the last years, clinicians and investigators have increasingly focused on additional strategies for improving outcomes and long-term prognosis of patients with RA. The effect of different strategies to reach clinical targets defined by composite measures is reflected in better mean levels of disease activity than in the past, and lower long-term consequences of poor control. However, with the improvement of disease outcome, questions regarding the usefulness of traditional measures arise. Overall, the expectations have been set higher than previously, in a way that the majority of patients should have the treatment boosted to achieve even better outcomes. However, treatment is not without side effects. This is especially applicable for patients with early disease whom would have a higher benefit of an earlier treatment, but whom may not all of them develop the worst forms of arthritis.

As we have seen during our research, most of the classic measures are not exempt of certain bias. Moreover, there are not studies evaluating the feasibility of implementing them in daily practice, especially in patients with early arthritis. The implementation of measures into routine care may be stimulated by the selection of simple and comprehensible tools. Different variations of DAS have been proposed, as a version including reduced joint counts or an alternative acute phase reactant. Even CDAI was developed as an alternative to SDAI without acute phase reactants to be more applicable in routine care. Although multiple attempts to improve these composite indexes have been done in the last years there is not unanimity

concerning the best index to be used. Research regarding successful implementation is limited and few implementation strategies have shown some value (van Hulst, Creemers et al. 2010). The heterogeneity and complexity of these measures, as well as their perceived expense in time and effort, can contribute to some reluctance by rheumatologists to perform them regularly and to change treatments accordingly. It is important to improve these measures for a successful use in routine care. New medications, especially sophisticated biologics, are used and led to outcomes previously unattainable, such as remission. However, these drugs are expensive, and this has resulted in pressure for physicians not only to use these drugs in selected patients for whom these drugs might be effective, but also pressure to document efficacy of these drugs in individual patients longitudinally. These new aspects have changed the clinical practice of rheumatology along with the need of objective disease measurements.

## **Treat to target in early and undifferentiated arthritis**

Adopting a tight control approach to the management of patients with RA yields superior clinical outcomes (Schoels, Knevel et al. 2010). While this approach has been successfully implemented in other chronic diseases, such as diabetes or hypercholesterolemia, its use in RA is less straightforward as there is not a simple “gold standard” measure for disease activity. A key component of the tight control approach is the availability of easily implemented measures that allow physicians to monitor disease activity. Recently, an international expert committee elaborated recommendations to enable the implementation of a ‘treat to target’ (T2T) approach into daily clinical practice (Smolen, Aletaha et al. 2010). Part of the recommendations is to establish routine visits to document treatment efficacy scheduled every three months. An increased frequency of visits is suggested during phases of higher disease activity, being the ideal target to reach remission. Remission is defined as the absence of signs and symptoms of significant inflammatory disease activity. This target of remission can be applied particularly to patients with newer onset RA, whereas in longstanding and refractory disease a low disease activity may be acceptable. Several randomized comparisons of T2T versus routine care have shown the benefits of targeted treatment in early RA (Grigor, Capell et al. 2004; Verstappen, Jacobs et al. 2007; van Tuyl, Lems et al. 2008; Soubrier, Lukas et al. 2011). Although the majority of studies reach similar conclusions, it is difficult to compare results owing to the heterogeneity of measures and clinical targets. These include composite scores like the DAS and DAS28, the CDAI and SDAI, the RAPID3, or single measures as swollen joint count.

Patients with early RA may have a higher likelihood of achieving remission because disease has not yet caused damage so they have higher chance of achieving normal function. What targets patients and clinicians should aim for remains to be debated. Comorbidities, such as pulmonary disease, or the risk of infection, may limit the therapy, and must be considered. Moreover, it is not known how feasible the implementation of T2T can be in real-world routine care, as collecting the variables to score composite indices in all visits is not always achievable. Only a small percentage of rheumatologist routinely perform joint counts (Pincus and Segurado 2006).

Several reports suggest improved mortality outcomes in RA (Bjornadal, Baecklund et al. 2002) not only due to a better use of classic treatments and the availability of newer effective therapies, such as the biologics, but because of the increasing use of suitable measures of disease activity to monitor patients with RA as well. The next reasonable step would be to use this efficacious approach in early and undifferentiated arthritis patients. Patients with early RA may have higher chances to reach better outcomes, and probably reach easily remission. Remission is now a more realistic target and it is necessary that clinicians have an accurate and reliable measure to define remission in patients with early arthritis. While certain composite indices—such as DAS and SDAI—are used in clinical trials, these complex instruments are very difficult to implement for busy rheumatology clinics. For this reason we need to implement more feasible indices that can be used in routine care to treat patients following a T2T strategy as earlier as possible, and here HUPI could be a useful tool.

## Future perspectives

Daily practice in rheumatology is characterized by heterogeneity of patients, mainly in treatment tolerance or indication and patient's expectations. In order to define treatment targets it is also important to have in perspective the perception of disease activity by the patient. In the last years, patient perspective has been emphasised at different levels. Some years ago the OMERACT 6 conference re-examined the core set of current outcomes measures for RA from the patient perspective. Whether the priorities and concerns of patients were being adequately incorporated into assessments of disease severity and progression was then explored (Kirwan, Heiberg et al. 2003). Although patient reported outcomes (PROs) are becoming more frequently used, this is done with great heterogeneity, as observed in a recent systemic literature review, aimed to assess the frequency of use of different PROs in published RA articles in the previous two years (Kalyoncu, Dougados et al. 2009).

A strategy to incorporate the patient perspective in clinical care could be to use composite indices based on PRO. For example, RAPID3, an index that includes only the three PRO measures in the Core Data Set—physical function, pain, and global estimate of disease activity. This index correlates significantly with the DAS in clinical trials (Pincus, Furer et al. 2011) and clinical care (Pincus, Swearingen et al. 2008) and can distinguish active from control treatments at levels similar to ACR or DAS criteria in different clinical trials (Pincus, Strand et al. 2003; Pincus, Chung et al. 2006). PROs not only can provide an easily implemented approach for assessment of patients with early and undifferentiated arthritis in usual care settings; they can be also useful to implement the T2T strategy (Castrejon and Pincus 2012). Moreover, PRO present additional advantages: the same observer—the patient—completes the quantitative information, what reduces variability; the patient, not the doctor, fills in all data needed to calculate the composite index, what improves visit efficiency; and finally, self-assessment helps the patient being prepared for the visit, what improves doctor-patient communication. The down side of using PROs is that some physicians consider them subjective and less valid than “objective” measures. Interestingly, the interpretation of patient global assessment is

more consistent and reliable than physician global assessment and joint counts (Aletaha, Machold et al. 2006). In addition, it is well established that patient and physician do not score disease activity similarly (Nicolau, Yogui et al. 2004; van Tuyl, Plass et al. 2008; Barton, Imboden et al. 2010; Khan, Spencer et al. 2012). In general, physicians tolerate higher values on the patient global scale than on the physician global scale. This suggests that physicians tend to assume that patients rate their disease activity higher than their physicians, what is in agreement with observed results in various studies of rheumatic diseases (van Tuyl, Plass et al. 2008; Barton, Imboden et al. 2010; Hudson, Impens et al. 2010; Castrejon, Yazici et al. 2013). Because of the absence of a valid “gold standard” for global disease activity, it is difficult to know whom—physicians or patients—are closer to the true global disease activity. Although, laboratory data can be seen as “objective” measures they can be normal in certain percentage of patients and they cannot capture the complete spectrum of the disease. Variables reported by the patient as pain, function, fatigue and others should also be included for a more complete understanding of the disease.

Finally, there is an increased pressure for registering disease activity from insurance companies for reimbursement of treatment (Hobbs and Cohen 2012) and from health systems for assessment of performance, in the form of quality of care indicators (van Hulst, Fransen et al. 2009); in all, these pressures will force us, the rheumatology community, to do our best in term of measuring.

## **PART IV: CONCLUSIONS**

We have gained good insight in the measurement of rheumatoid arthritis, concretely in the measurement of early and undifferentiated arthritis.

The results of our research lead us to the following conclusions:

1. DAS28-ESR and DAS28-CRP are not fully equivalent, the former yielding higher values than the latter, on average; this should be taken into account when using these indices, particularly in women and in patients with longstanding disease.
2. The cut-off points for DAS28-CRP in early arthritis should be lower than those described for DAS28-ESR in established disease: < 2.3 for remission, 2.3–3.8 for low disease activity, 3.8–4.9 for moderate disease activity, and > 4.9 for high disease activity; the cut-off value for minimal disease activity is 2.6.
3. Women with early rheumatoid arthritis have higher DAS28-ESR scores because of higher tender joint counts and erythrocyte sedimentation rate; this may represent a bias when assessing the response to treatment using the DAS28-ESR, which can be partially avoided by using DAS28-CRP.
4. There is no direct evidence on what index should be used to follow-up patients with undifferentiated peripheral inflammatory arthritis, or at which intervals these should be assessed.
5. In light of very limited evidence and mainly based on expert experience, disease activity should be monitored in undifferentiated arthritis; however, no specific instrument can be recommended; research is needed to evaluate the validity of the most common or new indices in undifferentiated arthritis.
6. We have developed and validated a new composite disease activity index based on the sum of 4 variables graded from 0 to 3: tender joint count, swollen joint count, patient global disease assessment, and erythrocyte sedimentation rate/C-reactive protein level. Hospital Universitario de La Princesa Index is feasible and sensitive to change, and could prove superior to previous indices in that it prevents bias arising from sex and missing data. This new index can be used in patients with early and undifferentiated arthritis.

## **PARTE IV: CONCLUSIONES**

Hemos adquirido buen conocimiento sobre la medición de la actividad de la enfermedad en la artritis reumatoide, y en concreto en la medida de artritis precoz e indiferenciada.

Los resultados de nuestra investigación nos han conducido a desarrollar las siguientes conclusiones:

1. El DAS28-VSG y el DAS28-PCR no son totalmente equivalentes, con el primero se obtienen en promedio valores más altos. Este aspecto se debe tener en cuenta, en particular en mujeres y en pacientes con enfermedad de larga duración.
2. Los puntos de corte para el DAS28-PCR en artritis precoz deberían ser más bajos que aquellos descritos para el DAS28-VSG: < 2,3 para remisión, 2,3–3,8 para actividad baja, 3,8–4,9 para actividad moderada, y > 4,9 para actividad alta; el punto de corte para actividad mínima de la enfermedad es 2,6.
3. Las mujeres con artritis reumatoide precoz tienen un DAS28-VSG más alto debido a una puntuación más alta en el número de articulaciones dolorosas y la velocidad de sedimentación globular. Esto puede representar un sesgo a la hora de evaluar la respuesta al tratamiento usando el DAS28-VSG, lo cual podría ser parcialmente evitado al usar DAS28-PCR.
4. No hay evidencia directa sobre qué índice debería ser empleado para seguir a los pacientes con artritis periférica inflamatoria indiferenciada, ni tampoco con qué frecuencia deberían ser evaluados.
5. Dada la evidencia limitada y basándonos en la experiencia de los expertos, la actividad de la enfermedad debería ser evaluada en la artritis indiferenciada. Sin embargo, no se puede recomendar ningún instrumento en concreto. Es necesario realizar más investigación sobre la validación de los instrumentos de medida más comunes en artritis indiferenciada.
6. Hemos desarrollado y validado un nuevo índice compuesto de actividad de la enfermedad basado en la suma de 4 variables con una puntuación que va de 0 a 3: el número de articulaciones dolorosas y tumefactas, la evaluación global de la enfermedad por el paciente, la velocidad de sedimentación globular y la proteína C-reactiva. El HUPI es factible y sensible al cambio, y podría ser superior a otros índices ya descritos a la hora de prevenir sesgos de género y de falta de datos. Este nuevo índice puede ser utilizado en pacientes con artritis precoz e indiferenciada.



## REFERENCE LIST

- Aletaha, D., G. Eberl, et al. (2004). "Attitudes to early rheumatoid arthritis: changing patterns. Results of a survey." *Ann Rheum Dis* **63**(10): 1269-1275.
- Aletaha, D., R. Landewe, et al. (2008). "Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations." *Arthritis Rheum* **59**(10): 1371-1377.
- Aletaha, D., K. P. Machold, et al. (2006). "The perception of rheumatoid arthritis core set measures by rheumatologists. Results of a survey." *Rheumatology (Oxford)* **45**(9): 1133-1139.
- Aletaha, D., V. P. Nell, et al. (2005). "Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score." *Arthritis Res Ther* **7**(4): R796-806.
- Aletaha, D., T. Neogi, et al. (2010). "2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative." *Ann Rheum Dis* **69**(9): 1580-1588.
- Aletaha, D. and J. Smolen (2005). "The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis." *Clin Exp Rheumatol* **23**(5 Suppl 39): S100-108.
- Aletaha, D., M. M. Ward, et al. (2005). "Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states." *Arthritis Rheum* **52**(9): 2625-2636.
- Anderson, J., L. Caplan, et al. (2012). "Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice." *Arthritis Care Res (Hoboken)* **64**(5): 640-647.
- Arnett, F. C., S. M. Edworthy, et al. (1988). "The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis." *Arthritis Rheum* **31**(3): 315-324.
- Bakker, M. F., J. W. Jacobs, et al. (2007). "Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility." *Ann Rheum Dis* **66** Suppl 3: iii56-60.
- Barton, J. L., J. Imboden, et al. (2010). "Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis." *Arthritis Care Res (Hoboken)* **62**(6): 857-864.
- Bjornadal, L., E. Baecklund, et al. (2002). "Decreasing mortality in patients with rheumatoid arthritis: results from a large population based cohort in Sweden, 1964-95." *J Rheumatol* **29**(5): 906-912.

- Boutry, N., E. Hachulla, et al. (2005). "MR imaging findings in hands in early rheumatoid arthritis: comparison with those in systemic lupus erythematosus and primary Sjogren syndrome." *Radiology* **236**(2): 593-600.
- Carbonell, J., T. Cobo, et al. (2008). "The incidence of rheumatoid arthritis in Spain: results from a nationwide primary care registry." *Rheumatology (Oxford)* **47**(7): 1088-1092.
- Carmona, L., J. Ballina, et al. (2001). "The burden of musculoskeletal diseases in the general population of Spain: results from a national survey." *Ann Rheum Dis* **60**(11): 1040-1045.
- Castrejon, I. and T. Pincus (2012). "Patient self-report outcomes to guide a treat-to-target strategy in clinical trials and usual clinical care of rheumatoid arthritis." *Clin Exp Rheumatol* **30 (Suppl. 73)**: S50-S55.
- Castrejon, I., Y. Yazici, et al. (2013). "Discordance of global estimates by patients and their physicians in usual care of many rheumatic diseases is associated with 5 MDHAQ scores not found on the HAQ." *Arthritis Care Res (Hoboken)*.
- de Bruin, A. F., J. P. Diederiks, et al. (1997). "Assessing the responsiveness of a functional status measure: the Sickness Impact Profile versus the SIP68." *J Clin Epidemiol* **50**(5): 529-540.
- Deodhar, S. D. (1989). "C-reactive protein: the best laboratory indicator available for monitoring disease activity." *Cleve Clin J Med* **56**(2): 126-130.
- Feldman, D. E., O. Schieir, et al. (2009). "RapiditY of rheumatology consultation for people in an early inflammatory arthritis cohort." *Ann Rheum Dis* **68**(11): 1790-1791.
- Felson, D. T., J. J. Anderson, et al. (1993). "The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials." *Arthritis Rheum* **36**(6): 729-740.
- Fransen, J., K. Visser, et al. (2010). "Validity of the disease activity score in undifferentiated arthritis." *Arthritis Care Res (Hoboken)* **62**(10): 1392-1398.
- Fransen J, W. P. (2004). "Disease activity score using C-reactive protein: CRP may replace ESR in the assessment of RA disease activity. ." *Ann Rheum Dis*
- Gabay, C. and I. Kushner (1999). "Acute-phase proteins and other systemic responses to inflammation." *N Engl J Med* **340**(6): 448-454.
- Gossec, L., M. Dougados, et al. (2004). "Prognostic factors for remission in early rheumatoid arthritis: a multiparameter prospective study." *Ann Rheum Dis* **63**(6): 675-680.
- Grigor, C., H. Capell, et al. (2004). "Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial." *Lancet* **364**(9430): 263-269.
- Hanley, J. A. and B. J. McNeil (1982). "The meaning and use of the area under a receiver operating characteristic (ROC) curve." *Radiology* **143**(1): 29-36.
- Hobbs, K. F. and M. D. Cohen (2012). "Rheumatoid arthritis disease measurement: a new old idea." *Rheumatology (Oxford)* **51 Suppl 6**: vi21-27.
- Hudson, M., A. Impens, et al. (2010). "Discordance between patient and physician assessments of disease severity in systemic sclerosis." *J Rheumatol* **37**(11): 2307-2312.
- Inoue, E., H. Yamanaka, et al. (2007). "Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values." *Ann Rheum Dis* **66**(3): 407-409.

- Kalyoncu, U., M. Dougados, et al. (2009). "Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review." *Ann Rheum Dis* **68**(2): 183-190.
- Khan, N. A., H. J. Spencer, et al. (2012). "Determinants of discordance in patients' and physicians' rating of rheumatoid arthritis disease activity." *Arthritis Care Res (Hoboken)* **64**(2): 206-214.
- Khera, A., D. K. McGuire, et al. (2005). "Race and gender differences in C-reactive protein levels." *J Am Coll Cardiol* **46**(3): 464-469.
- Kirwan, J., T. Heiberg, et al. (2003). "Outcomes from the Patient Perspective Workshop at OMERACT 6." *J Rheumatol* **30**(4): 868-872.
- Kumar, K., E. Daley, et al. (2007). "Delay in presentation to primary care physicians is the main reason why patients with rheumatoid arthritis are seen late by rheumatologists." *Rheumatology (Oxford)* **46**(9): 1438-1440.
- Kushner, I. (1991). "C-reactive protein in rheumatology." *Arthritis Rheum* **34**(8): 1065-1068.
- Makinen, H., H. Kautiainen, et al. (2005). "Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis?" *Ann Rheum Dis* **64**(10): 1410-1413.
- Matsui, T., Y. Kuga, et al. (2007). "Disease Activity Score 28 (DAS28) using C-reactive protein underestimates disease activity and overestimates EULAR response criteria compared with DAS28 using erythrocyte sedimentation rate in a large observational cohort of rheumatoid arthritis patients in Japan." *Ann Rheum Dis* **66**(9): 1221-1226.
- Mewar, D., A. Coote, et al. (2006). "Independent associations of anti-cyclic citrullinated peptide antibodies and rheumatoid factor with radiographic severity of rheumatoid arthritis." *Arthritis Res Ther* **8**(4): R128.
- Nell, V. P., K. P. Machold, et al. (2004). "Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis." *Rheumatology (Oxford)* **43**(7): 906-914.
- Nell, V. P., K. P. Machold, et al. (2005). "Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis." *Ann Rheum Dis* **64**(12): 1731-1736.
- Neogi, T., D. Aletaha, et al. (2010). "The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: Phase 2 methodological report." *Arthritis Rheum* **62**(9): 2582-2591.
- Nicolau, G., M. M. Yogui, et al. (2004). "Sources of discrepancy in patient and physician global assessments of rheumatoid arthritis disease activity." *J Rheumatol* **31**(7): 1293-1296.
- Nishimura, K., D. Sugiyama, et al. (2007). "Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis." *Ann Intern Med* **146**(11): 797-808.
- Pincus, T., L. F. Callahan, et al. (1984). "Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years." *Arthritis Rheum* **27**(8): 864-872.
- Pincus, T., C. Chung, et al. (2006). "An index of patient reported outcomes (PRO-Index) discriminates effectively between active and control treatment in 4 clinical trials of adalimumab in rheumatoid arthritis." *J Rheumatol* **33**(11): 2146-2152.

- Pincus, T., V. Furer, et al. (2011). "RAPID3 (Routine Assessment of Patient Index Data 3) severity categories and response criteria: Similar results to DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) in the RAPID 1 (Rheumatoid Arthritis Prevention of Structural Damage) clinical trial of certolizumab pegol." *Arthritis Care Res (Hoboken)* **63**(8): 1142-1149.
- Pincus, T. and O. G. Segurado (2006). "Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count." *Ann Rheum Dis* **65**(6): 820-822.
- Pincus, T., V. Strand, et al. (2003). "An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from that of placebo as effectively as the American College of Rheumatology 20% response criteria (ACR20) or the Disease Activity Score (DAS) in a rheumatoid arthritis clinical trial." *Arthritis Rheum* **48**(3): 625-630.
- Pincus, T., C. J. Swearingen, et al. (2008). "RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories." *J Rheumatol* **35**(11): 2136-2147.
- Prevoo, M. L., M. A. van 't Hof, et al. (1995). "Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis." *Arthritis Rheum* **38**(1): 44-48.
- Prevoo, M. L., A. M. van Gestel, et al. (1996). "Remission in a prospective study of patients with rheumatoid arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score." *Br J Rheumatol* **35**(11): 1101-1105.
- Quinn, M. A., P. G. Conaghan, et al. (2001). "The therapeutic approach of early intervention for rheumatoid arthritis: what is the evidence?" *Rheumatology (Oxford)* **40**(11): 1211-1220.
- Raza, K., F. Falciani, et al. (2005). "Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin." *Arthritis Res Ther* **7**(4): R784-795.
- Rose, H. M., C. Ragan, et al. (1948). "Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis." *Proc Soc Exp Biol Med* **68**(1): 1-6.
- Schoels, M., R. Knevel, et al. (2010). "Evidence for treating rheumatoid arthritis to target: results of a systematic literature search." *Ann Rheum Dis* **69**(4): 638-643.
- Scott, D. L., F. Wolfe, et al. (2010). "Rheumatoid arthritis." *Lancet* **376**(9746): 1094-1108.
- Smolen, J. S., D. Aletaha, et al. (2010). "Treating rheumatoid arthritis to target: recommendations of an international task force." *Ann Rheum Dis* **69**(4): 631-637.
- Smolen, J. S., F. C. Breedveld, et al. (2003). "A simplified disease activity index for rheumatoid arthritis for use in clinical practice." *Rheumatology (Oxford)* **42**(2): 244-257.
- Smolen, J. S., T. Sokka, et al. (2003). "A proposed treatment algorithm for rheumatoid arthritis: aggressive therapy, methotrexate, and quantitative measures." *Clin Exp Rheumatol* **21**(5 Suppl 31): S209-210.

- Soubrier, M., C. Lukas, et al. (2011). "Disease activity score-driven therapy versus routine care in patients with recent-onset active rheumatoid arthritis: data from the GUEPARD trial and ESPOIR cohort." Ann Rheum Dis.
- Thompson, D., A. Milford-Ward, et al. (1992). "The value of acute phase protein measurements in clinical practice." Ann Clin Biochem **29 ( Pt 2)**: 123-131.
- Tugwell, P. and M. Boers (1993). "Developing consensus on preliminary core efficacy endpoints for rheumatoid arthritis clinical trials. OMERACT Committee." J Rheumatol **20(3)**: 555-556.
- Tugwell, P. and M. Boers (1993). "OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction." J Rheumatol **20(3)**: 528-530.
- van der Heijde, D. M., M. van 't Hof, et al. (1993). "Development of a disease activity score based on judgment in clinical practice by rheumatologists." J Rheumatol **20(3)**: 579-581.
- van der Heijde, D. M., M. A. van 't Hof, et al. (1990). "Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score." Ann Rheum Dis **49(11)**: 916-920.
- van Gestel, A. M., C. J. Haagsma, et al. (1998). "Validation of rheumatoid arthritis improvement criteria that include simplified joint counts." Arthritis Rheum **41(10)**: 1845-1850.
- van Hulst, L. T., M. C. Creemers, et al. (2010). "How to improve DAS28 use in daily clinical practice?--a pilot study of a nurse-led intervention." Rheumatology (Oxford) **49(4)**: 741-748.
- van Hulst, L. T., J. Fransen, et al. (2009). "Development of quality indicators for monitoring of the disease course in rheumatoid arthritis." Ann Rheum Dis **68(12)**: 1805-1810.
- van Nies, J. A., A. Krabben, et al. (2013). "What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review." Ann Rheum Dis.
- van Tuyl, L. H., W. F. Lems, et al. (2008). "Tight control and intensified COBRA combination treatment in early rheumatoid arthritis: 90% remission in a pilot trial." Ann Rheum Dis **67(11)**: 1574-1577.
- van Tuyl, L. H., A. M. Plass, et al. (2008). "Discordant perspectives of rheumatologists and patients on COBRA combination therapy in rheumatoid arthritis." Rheumatology (Oxford) **47(10)**: 1571-1576.
- Verstappen, S. M., J. W. Jacobs, et al. (2007). "Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial)." Ann Rheum Dis **66(11)**: 1443-1449.
- Wells, G. A., M. Boers, et al. (2005). "Minimal disease activity for rheumatoid arthritis: a preliminary definition." J Rheumatol **32(10)**: 2016-2024.
- Welsing, P. M., J. Fransen, et al. (2005). "Is the disease course of rheumatoid arthritis becoming milder? Time trends since 1985 in an inception cohort of early rheumatoid arthritis." Arthritis Rheum **52(9)**: 2616-2624.
- Wolfe, F., J. J. Rasker, et al. (2007). "Minimal disease activity, remission, and the long-term outcomes of rheumatoid arthritis." Arthritis Rheum **57(6)**: 935-942.

## ABBREVIATION LIST

ACR: American College of Rheumatology  
Anti-CCP: Anti-citrullinated protein antibodies  
APR: Acute Phase Reactants  
AUC: Area Under the Curve  
CDAI: Clinical Disease Activity Index  
CRP: C-reactive protein  
DAS: Disease Activity Score  
DAS28: Disease Activity Score 28  
DMARDs: Disease-Modifying Antirheumatic Drugs  
EAC: Early Arthritis Clinic  
ESR: Erythrocyte Sedimentation Rate  
EULAR: European League Against Rheumatism  
HAQ: Health Assessment Questionnaire  
HUPI: Hospital Universitario de La Princesa Index  
MDA: Minimal disease activity  
MRI: Magnetic Resonance Imaging  
NSAIDs: Nonsteroidal anti-inflammatory drugs  
OMERACT: Outcome Measures in Rheumatoid Arthritis Clinical Trials  
PGA: patient's global assessments of disease activity  
PhGA: physician global assessment of disease activity  
PEARL: Princess Early Arthritis Register Longitudinal  
PROs: patient reported outcomes  
RA: Rheumatoid Arthritis  
RCT: Randomized Clinical Trials  
RF: Rheumatoid Factor  
ROC curves: Receiver Operating Characteristic curves  
SDAI: Simplified Disease Activity Index  
Se: Sensitivity  
SJC: Swollen Joint Count  
Sp: Specificity  
TJC: Tender Joint Count  
T2T: Treat To Target  
UA: Undifferentiated Arthritis



UPIA: undifferentiated peripheral inflammatory arthritis

VAS: Visual Analogue Scale

## **PART V: ADDENDUM**

### **Role of the funding sources**

The rheumatologists participating in the Early Arthritis Clinic at Hospital Universitario La Princesa were responsible for the data collection and study design. The authors are responsible for the data analysis, interpretation of data and preparing the manuscripts.

The 3E Initiative was supported by Abbott with an unrestricted educational grant.

### **Curriculum Vitae**

Dr. Isabel Castrejón graduated from Alcala de Henares Medical University in 2001. She was trained as Rheumatologist at Hospital Universitario La Princesa in Madrid (2007), and worked as Rheumatologist at Hospital La Princesa and Hospital 12 de Octubre in Madrid before moving to New York. She has been working in the Division of Rheumatology at NYU Hospital for Joint Diseases since 2010 where she has been participating in a research program on outcomes in rheumatic diseases under the supervision of Professor Theodore Pincus (Director of Outcomes Research).

She participated as clinical fellow at the Division of Pediatric Rheumatology in 2007 (Chief of the Division Dr. Thomas JA Lehman. Hospital for Special Surgery, NYC) and at the Bellevue Hospital and the Seligman Center for Advanced Therapeutics in 2009, NYC, where she participated in the Lupus Clinic and validated the MDHAQ questionnaire in lupus patients (Dr. Anca Askanase and Professor Theodore Pincus)

She was part of a taskforce for the development of a EULAR (European League Against Rheumatism) toolbox online for patient report outcomes and she has been an active member of the EMEUNET (Emerging EULAR Network) as subgroup leader for the Peer (mentoring and research collaboration) group. She has also undergone different Epidemiology and Biostatistics programs, at the University of Michigan (Ann Arbor, USA), at the University of Aberdeen, School of Medicine and Dentistry (UK) and at Hunter College, Cuny School of Public Health (NYC). Her areas of interest include outcomes research (validation of questionnaires and disease activity indices), rheumatoid arthritis, lupus, and clinical epidemiology.

She has presented more than 50 abstracts, including oral presentations, to international conferences and co-authored around 35 papers on a variety of topics, with an emphasis on patient report outcomes and disease activity measures.

Dr. Castrejon is an ACR (American College of Rheumatology) International Fellow Member and holds a full registration with a license to practice and entry on the Specialist Register as Rheumatologist in the General Medical Council (UK).

## List of publications

1. **Castrejon I**, Yazici Y, Samuels J, Luta G, Pincus T. Discordance of global estimates by patients and their physicians in usual care of many rheumatic diseases is associated with 5 MDHAQ scores not found on the HAQ. *Arthritis Care Res (Hoboken)*. 2013 Dec 2. doi: 10.1002/acr.22237. [Epub ahead of print]
2. **Castrejon I**, Gibson KA, Pincus T. Efficacy and Safety of Methotrexate in Combination with Other Non-biologic Disease-modifying Antirheumatic Drugs (DMARDs) in Treatment of Rheumatoid Arthritis. *Bull Hosp Jt Dis (2013)* 2013;71 Suppl 1:20-8.
3. Pincus T, **Castrejon I**. Evidence that the Strategy is More Important than the Agent to Treat Rheumatoid Arthritis - Data from Clinical Trials of Combinations of Non-Biologic DMARDs, with Protocol-Driven Intensification of Therapy for Tight Control or Treat-to-Target. *Bull Hosp Jt Dis (2013)* 2013;71 Suppl 1:33-40
4. Pincus T, Gibson KA, **Castrejon I**. Update on methotrexate as the anchor drug for rheumatoid arthritis. *Bull Hosp Jt Dis (2013)* 2013;71 Suppl 1:9-19.
5. **Castrejon I**, Pincus T, Soubrier M, Lin Y, Rat A, Combe B, et al. GUEPARD treat-to-target strategy is significantly more efficacious than ESPOIR routine care in early rheumatoid arthritis according to patient-reported outcomes and physician global estimate. *Rheumatology (Oxford)* 2013 Oct;52:1980-7
6. Pincus T, **Castrejon I**, Yazici Y. Documenting the value of care for rheumatoid arthritis, analogous to hypertension, diabetes, and hyperlipidemia: is control of individual patient self-report measures of global estimate and physical function more valuable than laboratory tests, radiographs, indices, or remission criteria? *J Rheumatol* 2013 Sep;40:1469-74
7. **Castrejon I**, Bergman MJ, Pincus T. MDHAQ/RAPID3 to recognize improvement over 2 months in usual care of patients with osteoarthritis, systemic lupus erythematosus, spondyloarthropathy and gout. *J Clin Rheumatol*. 2013 Jun;19(4):169-74.
8. Pincus T, Sokka T, **Castrejon I**, Cutolo M. Decline of mean initial prednisone dose from 10.3 to 3.6 mg/day to treat rheumatoid arthritis between 1980 and 2004 in one clinical setting, with long-term effectiveness of doses less than 5 mg/day. *Arthritis Care Res (Hoboken)*. 2013 May;65(5):729-36
9. **Castrejon I**, Carmona L, Ortiz AM, Belmonte MA, et al. Development and validation of a new disease activity index as a numerical sum of 4 variables in patients with early arthritis. *Arthritis Care Res (Hoboken)*. 2013 Apr;65:518-25
10. **Castrejon I**, Dougados M, Combe B, Guillemin F, Fautrel B, Pincus T. Can remission in rheumatoid arthritis be assessed without laboratory tests or a formal joint count?

- Possible Remission Criteria Based on a Self-report RAPID3 Score and Careful Joint Examination in the ESPOIR Cohort. *The Journal of Rheumatology* 2013 Apr; 40:386-93
11. **Castrejon I**, Yazici Y, Pincus T. Patient self-report RADAI (Rheumatoid Arthritis Disease Activity Index) joint counts on an MDHAQ (Multidimensional Health Assessment Questionnaire) in usual care of consecutive patients with rheumatic diseases other than rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2013 Feb; 65:288-93.
  12. Pincus T, **Castrejon I**, Yazici Y. Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis. *Ann Intern Med.* 2012 Aug 21; 157(4):299
  13. **Castrejon I**, McCollum L, Tanriover MD, Pincus T. Importance of patient history and physical examination in rheumatoid arthritis compared to other chronic diseases: Results of a physician survey. *Arthritis Care Res (Hoboken)*. 2012 Aug; 64(8):1250-1255.
  14. **Castrejon I**, Pincus T. Patient self-report outcomes to guide a treat-to-target strategy in clinical trials and usual clinical care of rheumatoid arthritis. *Clin Exp Rheumatol* 2012 Jul-Aug;30 (Suppl. 73):S50-S5.
  15. Pincus T, **Castrejon I**, Bergman M, Yazici Y. Treat-to-target: not as simple as it appears. *Clin Exp Rheumatol* 2012 Jul-Aug; 30 (Suppl. 73):S10-S20.
  16. Pincus T, Yazici Y, **Castrejon I**. Pragmatic and Scientific Advantages of MDHAQ/ RAPID3 Completion by All Patients at All Visits in Routine Clinical Care. *Bull NYU Hosp Jt Dis* 2012;70 Suppl 1:30-6.
  17. Pincus T, **Castrejon I**. Effective Initial and Long-Term Prednisone in Doses of Less Than 5 Mg/Day to Treat Rheumatoid Arthritis - Documentation Using a Patient Self-Report Multidimensional Health Assessment Questionnaire (MDHAQ). *Bull NYU Hosp Jt Dis* 2012;70 Suppl 1:14-20.
  18. Pincus T, Skummer PT, Grisanti MT, **Castrejon I**, Yazici Y. MDHAQ/RAPID3 Can Provide a Roadmap or Agenda for all Rheumatology Visits When the Entire MDHAQ is Completed at All Patient Visits and Reviewed by the Doctor Before the Encounter. *Bull NYU Hosp Jt Dis* 2012;70:177-86.
  19. Pincus T, **Castrejon I**. An Evidence-Based Medical Visit for Patients with Rheumatoid Arthritis Based on Standard, Quantitative Scientific Data from a Patient MDHAQ and Physician Report. *Bull NYU Hosp Jt Dis.* 2012; 70(2):73-94.
  20. **Castrejon I**. [Transitional care programs for patients with rheumatic diseases: review of the literature]. *Reumatol Clin.* 2012 Jan-Feb; 8(1):20-26.
  21. **Castrejon I**, Shum K, Tseng CE, Askanase A. Association between myasthenia gravis and systemic lupus erythematosus: three case reports and review of the literature. *Scand J Rheumatol.* 2011 Nov; 40(6):486-90.
  22. Pincus T, **Castrejon I**, Sokka T. Long-term prednisone in doses of less than 5 mg/day for treatment of rheumatoid arthritis: personal experience over 25 years. *Clin Exp Rheumatol.* 2011 Sep-Oct; 29(5 Suppl 68):S130-138.
  23. Askanase AD, **Castrejon I**, Pincus T. Quantitative data for care of patients with systemic lupus erythematosus in usual clinical settings: a patient Multidimensional Health Assessment Questionnaire and physician estimate of noninflammatory symptoms. *J Rheumatol.* 2011 Jul; 38(7):1309-1316.

24. **Castrejon I**, Silva-Fernandez L, Bombardier C, Carmona L. Clinical composite measures of disease activity for diagnosis and followup of undifferentiated peripheral inflammatory arthritis: a systematic review. *J Rheumatol Suppl.* 2011 Mar; 87:48-53.
25. Silva-Fernandez L, **Castrejon I**, Bombardier C, Carmona L. Diagnostic and prognostic value of genetics in undifferentiated peripheral inflammatory arthritis: a systematic review. *J Rheumatol Suppl.* 2011 Mar; 87:38-44.
26. Machado P, **Castrejon I**, Katchamart W, Koevoets R, Kuriya B, Schoels M, et al. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Ann Rheum Dis.* 2011 Jan; 70(1):15-24.
27. Pincus T, **Castrejon I**. MDHAQ/RAPID3 Scores: Quantitative Patient History Data in a Standardized "Scientific" Format for Optimal Assessment of Patient Status and Quality of Care in Rheumatic Diseases. *Bull NYU Hosp Jt Dis.* 2011; 69(3):201-214.
28. **Castrejon I**, Pincus T. Differences in Treat-to-Target in Patients with Rheumatoid Arthritis versus Hypertension and Diabetes - Consequences for Clinical Care. *Bull NYU Hosp Jt Dis.* 2011; 69(2):104-110.
29. **Castrejon I**, Ortiz AM, Toledano E, Castañeda S, Garcia-Vadillo A, Patiño E, et al. Estimated cutoff points for the 28-joint disease activity score based on C-reactive protein in a longitudinal register of early arthritis. *J Rheumatol.* 2010 Jul; 37(7):1439-1443.
30. Ibanez M, Ortiz AM, **Castrejon I**, Garcia-Vadillo JA, Carvajal I, Castañeda S, et al. A rational use of glucocorticoids in patients with early arthritis has a minimal impact on bone mass. *Arthritis Res Ther.* 2010 Mar; 12(2):R50.
31. **Castrejon Fernandez I**, Sanz Sanz J. [Conventional Radiology: Total BASRI and SASSS]. *Reumatol Clin.* 2010 Mar; 6S1:33-36.
32. **Castrejon Fernandez I**, Martinez-Lopez JA, Ortiz Garcia AM, Carmona Ortells L, Garcia-Vicuña R, Gonzalez-Alvaro I. [Influence of gender on treatment response in a cohort of patients with early rheumatoid arthritis in the area 2 of Madrid]. *Reumatol Clin.* 2010 May-Jun; 6(3):134-140.
33. **Castrejon I**, Ortiz AM, Garcia-Vicuña R, Lopez-Bote JP, Humbria A, Carmona L, et al. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? *Clin Exp Rheumatol.* 2008 Sep-Oct; 26(5):769-775.
34. **Castrejon I**, Ibañez M, Vicente E, Luis Steegmann J, Castañeda S. [Relapsing polychondritis associated with a lymphoplasmocytic lymphoma and erythema nodosum]. *Reumatol Clin.* 2007 Jan; 3(1):45-47.